现代数学基础

4

# 高等概率论及其应用

■ 胡迪鹤 著



高等教育出版社
HIGHER EDUCATION PRESS

**内容简介**

本书是在初等概率论、测度论和泛函分析初步的基础上展开的。全书共分三大部分：一、高等概率的基本概念与工具，诸如随机元（含特例随机变量）及其分布，随机元的特征泛函，各种收敛性（含依概率收敛、概率为 1 地收敛、$L^p$ 收敛、完全收敛、淡收敛、局部弱收敛及弱收敛等）；二、概率极限理论，包括大数定律，中心极限定理，重对数律，不变原理，无穷可分律的理论及其应用等；三、随机过程论，包括可数状态离散时间的马尔可夫链，可数状态连续时间的马尔可夫过程，随机环境中马尔可夫链，鞅论等。在每章的最后，附有习题与应用。

本书是研究生的教学用书，也可供概率论的理论研究工作者、概率论与数理统计的应用研究工作者参考。

**图书在版编目（CIP）数据**

高等概率论及其应用 / 胡迪鹤著. —北京：高等教育出版社，2008.6

ISBN 978 - 7 - 04 - 022617 - 1

Ⅰ. 高 … Ⅱ. 胡 … Ⅲ. 概率论 Ⅳ. O211

中国版本图书馆 CIP 数据核字（2008）第 041784 号

| | | | |
|---|---|---|---|
| 出版发行 | 高等教育出版社 | 购书热线 | 010 - 58581118 |
| 社　　址 | 北京市西城区德外大街 4 号 | 免费咨询 | 800 - 810 - 0598 |
| 邮政编码 | 100120 | 网　　址 | http://www.hep.edu.cn |
| 总　　机 | 010 - 58581000 | | http://www.hep.com.cn |
| | | 网上订购 | http://www.landraco.com |
| 经　　销 | 蓝色畅想图书发行有限公司 | | http://www.landraco.com.cn |
| 印　　刷 | 北京奥鑫印刷厂 | 畅想教育 | http://www.widedu.com |
| 开　　本 | 787×1092　1/16 | 版　　次 | 2008 年 6 月第 1 版 |
| 印　　张 | 29.25 | 印　　次 | 2008 年 6 月第 1 次印刷 |
| 字　　数 | 550 000 | 定　　价 | 59.00 元 |



**物料号　22617 - 00**

# 前 言

本书是基于初等概率论、测度论和泛函分析初步而创作的。全书共分三大部分: 一、高等概率论的基本概念和工具, 诸如随机元及其分布, 随机元的特征泛函, 概率收敛, 概率为 1 地收敛, $L^p$ 收敛, 完全收敛, 淡收敛, 局部弱收敛, 弱收敛等。二、概率极限理论, 包括大数定律, 中心极限定理, 重对数律, 不变原理, 无穷可分律的理论及应用等。三、随机过程论, 包括可数状态离散时间的 Markov 链, 可数状态连续时间的 Markov 过程, 随机环境中的 Markov 链, 鞅论等。

本书的创作意图: 为硕士研究生和博士研究生提供教材参考书; 为涉及概率论的理论研究工作者和概率论与数理统计的应用研究工作者及将来有志于概率论与数理统计的学者提供一本参考书。作为硕士研究生的教材, 书中标有星号的章、节可略去不讲, 因为本书想照顾更多读者, 篇幅较大。

本书每章最后, 均附有习题及应用供教师及读者参考。书前附有常用符号表以备查用。

本书内容涉及面广, 体系繁复, 不妥之处, 包括取材、体系安排、表述及论证, 请读者不吝指教。

感谢国家自然科学基金委员会的资助 (项目批准号为: 10371092 和 10771185)。

作者 胡迪鹤

2007 年 7 月

# 常用符号表

$\mathbf{R}^d$: 表示 $d$ 维欧氏空间, $\mathbf{R}=\mathbf{R}^1, \mathbf{R}_+=[0,\infty)$.

$\mathscr{B}(\mathscr{G})$: 表示拓扑空间 $\mathscr{G}$ 中的 Borel $\sigma$ 代数, 即 $\mathscr{G}$ 中一切开集所产生的 $\sigma$ 代数.

$\mathscr{B}^d \triangleq \mathscr{B}(\mathbf{R}^d)$: 表示 $\mathbf{R}^d$ 中一切开集所产生的 $\sigma$ 代数, 即 Borel $\sigma$ 代数.

$b\mathscr{B}^d$: 表示一切有界 $\mathscr{B}^d$ 可测的实值函数.

$C([0,1])$: 表示 $[0,1]$ 上的一切实值连续函数.

$f \in \mathscr{F}_1/\mathscr{F}_2$: 表示 $f$ 关于 $\sigma$ 代数 $\mathscr{F}_1, \mathscr{F}_2$ 可测, 即 $f^{-1}(\mathscr{F}_2) \subset \mathscr{F}_1$. 特别地, $f \in \mathscr{F}_1$ 表示 $f^{-1}(\mathscr{B}^1) \subset \mathscr{F}_1$.

$a \wedge b$: 表示取 $a$ 与 $b$ 中最小者, 即 $a \wedge b = \min(a,b)$.

$a \vee b$: 表示取 $a$ 与 $b$ 中最大者, 即 $a \vee b = \max(a,b)$.

$\mathbf{Z}$: (如不特别声明) 表示整数集, $\mathbf{Z}_+$ 表示非负整数集.

$\boldsymbol{E}$ (斜黑体): 表示期望算子.

var: 表示方差.

cov: 表示协方差.

$(\Omega, \mathscr{F}, P)$: 表示概率空间.

$\bigvee\limits_{n\geqslant 0} \mathscr{F}_n$: 表示 $\sigma$ 代数列 $\{\mathscr{F}_n, n \geqslant 0\}$ 取并后以它 (未必是 $\sigma$ 代数) 所产生成的 $\sigma$ 代数.

设 $X=\{X_n, n\geqslant 0\}$ 是一实值随机过程, 则 $\|X_n\|^p$: 表示 $X_n$ 的 $L^p$ 范数, 即 $\boldsymbol{E}(|X_n|^p)$; $\|X\|^p$ 表示 $\sup\limits_{n\geqslant 0} \boldsymbol{E}(|X_n|^p)$.

R. V. : 表示随机变量.

d. f. : 表示分布函数.

d′. f. : 表示准分布函数.

c. f. : 表示特征函数.

i. d. : 表示无穷可分的.

$\xrightarrow{\mathrm{w}}$ (⟹): 表示关于测度的弱收敛.

$\overset{\mathrm{lw}}{\Longrightarrow}$: 表示关于测度的局部弱收敛.

$\overset{\mathrm{v}}{\Longrightarrow}$: 表示关于测度的淡收敛.

$\xrightarrow{P}$: 表示概率收敛.

$\xrightarrow{\mathrm{a.s.}}$: 表示概率为 1 地收敛.

$\xrightarrow{\mathrm{a.un.}}$: 表示几乎一致收敛.

$\xrightarrow{L^p}$: 表示 $L^p$ 收敛.

$\xrightarrow{\mathrm{com.}}$: 表示完全收敛.

$\xrightarrow{\mathrm{w}}$: 表示函数的弱收敛.

$\xrightarrow{\mathrm{lw}}$: 表示函数的局部弱收敛.

# 目　　录

# 第一章　距离空间中的测度

## §1　单调类定理

本书是基于测度论之上而写的, 但是, 为了读者的方便和本书的特殊需要, 对抽象空间中的测度的基本概念及性质, 对距离空间中的测度, 特别是距离空间中的 Hausdorff 测度和 $N$ 维欧氏空间中的 Lebesgue – Stieltjes 测度, 仍给以适量的论证.

关于集合的运算, 我们沿用习惯的称呼与符号. 例如, 给定集合 $\Omega$, 对其中任意两个子集 $A$ 和 $B$, 记为 $A \subset \Omega, B \subset \Omega$. 而 $A \bigcup B, A \bigcap B$ 和 $A - B$ 分别表示 $A$ 和 $B$ 之并、交与差. $A \triangle B$ 表示 $A$ 和 $B$ 的对称差, 即 $A \triangle B = (A - B) \bigcup (B - A)$. 类似地, 若 $\Gamma$ 是任一指标集, 且对任何 $r \in \Gamma, A_r \subset \Omega$, 仍用 $\bigcup\limits_{r \in \Gamma} A_r$ 和 $\bigcap\limits_{r \in \Gamma} A_r$ 分别表示 $\{A_r, r \in \Gamma\}$ 之并和交. $A^c = \Omega - A$ 表示 $A$ 之补集. $\varnothing$ 表示空集. $\mathbf{R}^N$ 表示 $N$ 维欧氏空间, $\mathbf{R} = \mathbf{R}^1, \overline{\mathbf{R}} = [-\infty, \infty]$. 对 $a, b \in \mathbf{R}, a \vee b \stackrel{\text{def.}}{=\!=} \max\{a, b\}, a \wedge b \stackrel{\text{def.}}{=\!=} \min\{a, b\}$.

**定义 1.1**　设 $\Omega$ 是任一给定的集合, $\mathfrak{M}$ 是 $\Omega$ 中一族子集, 称之为 $\Omega$ 的一个集合系.

(1) 称集合系 $\mathfrak{M}$ 是 $\pi$ **系**, 如果

$$\{A_1, \cdots, A_n\} \subset \mathfrak{M} \Rightarrow \bigcap_{i=1}^{n} A_i \in \mathfrak{M}.$$

(2) 称集合系 $\mathfrak{M}$ 是 $d$ **系**, 如果

(i) $\Omega \in \mathfrak{M}$;

(ii) $A, B \in \mathfrak{M}, A \subset B \Rightarrow B - A \in \mathfrak{M}$;

(iii) $A_n \in \mathfrak{M}, A_n \subset A_{n+1} (n \geqslant 1) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$.

(3) 称集合系 $\mathfrak{M}$ 是 $\sigma$ **代数**, 如果

(i) $\Omega \in \mathfrak{M}$;

(ii) $A, B \in \mathfrak{M} \Rightarrow A - B \in \mathfrak{M}$;

(iii) $A_n \in \mathfrak{M} (n \geqslant 1) \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$.

显然, 若 $\mathfrak{M}$ 是 $\sigma$ 代数, 则 $\varnothing \in \mathfrak{M}$, 而且当 $A_n \in \mathfrak{M} (n \geqslant 1)$ 时总有 $A_n^{\mathrm{c}} \in \mathfrak{M} (n \geqslant 1)$, 且 $\bigcap_{n=1}^{\infty} A_n = \left( \bigcup_{n=1}^{\infty} A_n^{\mathrm{c}} \right)^{\mathrm{c}} \in \mathfrak{M}$.

(4) 含 $\mathfrak{M}$ 的最小 $d$ 系 (或 $\sigma$ 代数) 称为由 $\mathfrak{M}$ 所产生的 $d$ 系 (或 $\sigma$ 代数), 记之为 $d(\mathfrak{M})$ (或 $\sigma(\mathfrak{M})$).

显然, $d(\mathfrak{M})$ (或 $\sigma(\mathfrak{M})$) 是唯一存在的.

**定理 1.1** 集合系 $\mathfrak{M}$ 是 $\sigma$ 代数的充分必要条件是: $\mathfrak{M}$ 既是 $\pi$ 系又是 $d$ 系.

**证** 必要性是显然的.

充分性. 由于 $\mathfrak{M}$ 既是 $\pi$ 系又是 $d$ 系, 所以欲证它是 $\sigma$ 代数, 只需证明它关于可数并运算封闭即可. 事实上, 由 $\mathfrak{M}$ 既是 $\pi$ 系又是 $d$ 系推知

$$A \in \mathfrak{M} \Rightarrow A^{\mathrm{c}} = \Omega - A \in \mathfrak{M}. \tag{1.1}$$

由 (1.1) 及 $\mathfrak{M}$ 是 $\pi$ 系推知

$$\{A_1, \cdots, A_n\} \subset \mathfrak{M} \Rightarrow \bigcup_{i=1}^{n} A_i = \left( \bigcap_{i=1}^{n} A_i^{\mathrm{c}} \right)^{\mathrm{c}} \in \mathfrak{M}. \tag{1.2}$$

由 (1.2) 式及 $d$ 系的第 (iii) 条性质可得

$$A_n \in \mathfrak{M} (n \geqslant 1) \Rightarrow \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \left( \bigcup_{i=1}^{n} A_i \right) \in \mathfrak{M}.$$

定理 1.1 得证.

**定理 1.2** 设 $\mathfrak{M}$ 是 $\Omega$ 上的一个 $\pi$ 系, 则 $d(\mathfrak{M}) = \sigma(\mathfrak{M})$.

**证** 由于 $\sigma(\mathfrak{M})$ 是包含 $\mathfrak{M}$ 的一个 $d$ 系, 所以 $\sigma(\mathfrak{M}) \supset d(\mathfrak{M})$. 因此, 为证定理, 只需证明 $d(\mathfrak{M})$ 是 $\sigma$ 代数. 根据定理 1.1, 则只需证明 $d(\mathfrak{M})$ 是一个 $\pi$ 系即可. 事实上, 若令

$$\mathscr{D}_1 = \{B \in d(\mathfrak{M}) : B \bigcap A \in d(\mathfrak{M}) \text{ 对一切 } A \in \mathfrak{M} \text{ 成立}\},$$

要证 $\mathscr{D}_1 = d(\mathfrak{M})$. 由于 $\mathfrak{M}$ 是 $\pi$ 系, 所以 $\mathscr{D}_1 \supset \mathfrak{M}$. 若能证 $\mathscr{D}_1$ 是一个 $d$ 系, 则 $\mathscr{D}_1 \supset d(\mathfrak{M})$, 从而 $\mathscr{D}_1 = d(\mathfrak{M})$.

(i) 显然, $\Omega \in \mathscr{D}_1$.

(ii) 设 $B_1, B_2 \in \mathscr{D}_1, B_1 \subset B_2$, 则对任何 $A \in \mathfrak{M}$, 由 $\mathscr{D}_1$ 的定义有 $B_i \bigcap A \in d(\mathfrak{M})(i=1,2)$. 又因 $B_1 \bigcap A \subset B_2 \bigcap A$, 所以由 $d$ 系的性质 (ii) 有

$$(B_2 - B_1)\bigcap A = ((B_2 \bigcap A) - (B_1 \bigcap A)) \in d(\mathfrak{M}),$$

此即 $B_2 - B_1 \in \mathscr{D}_1$.

(iii) 设 $B_n \subset B_{n+1}, B_n \in \mathscr{D}_1 (n \geqslant 1)$, 则对任何 $A \in \mathfrak{M}$, 有 $(B_n \bigcap A) \in d(\mathfrak{M}), (B_n \bigcap A) \subset (B_{n+1} \bigcap A)$, 所以由 $d$ 系之性质 (iii) 有

$$\left(\bigcup_{n=1}^{\infty} B_n\right)\bigcap A = \bigcup_{n=1}^{\infty} (B_n \bigcap A) \in d(\mathfrak{M}),$$

此即

$$\bigcup_{n=1}^{\infty} B_n \in \mathscr{D}_1.$$

这就证明了 $\mathscr{D}_1$ 是一个 $d$ 系, 所以 $\mathscr{D}_1 = d(\mathfrak{M})$.

再令

$$\mathscr{D}_2 = \{B \in d(\mathfrak{M}) : (B \bigcap A) \in d(\mathfrak{M}) \text{ 对一切 } A \in d(\mathfrak{M}) \text{ 成立}\},$$

要证 $\mathscr{D}_2 = d(\mathfrak{M})$, 即 $d(\mathfrak{M})$ 是 $\pi$ 系. 事实上, 由 $\mathscr{D}_1 = d(\mathfrak{M})$ 知: $\mathscr{D}_2 \supset \mathfrak{M}$. 仿 $\mathscr{D}_1$, 可证 $\mathscr{D}_2$ 也是 $d$ 系, 所以 $\mathscr{D}_2 = d(\mathfrak{M})$. 定理证毕.

**定义 1.2** 设 $\Omega$ 是任一给定的集合, $\mathscr{F}$ 是 $\Omega$ 上的一个 $\sigma$ 代数, 即 $(\Omega, \mathscr{F})$ 是一个可测空间. $f : \Omega \mapsto \mathbf{R}$, 称 $f$ 是 $\mathscr{F}$ 可测函数, 如果对任何 $x \in \mathbf{R}$, 都有 $\{\omega \in \Omega : f(\omega) \leqslant x\} \in \mathscr{F}$. 特别地, 若 $f$ 可表示为 $f = \sum_{i=1}^{n} c_i \mathbf{1}_{A_i}$, 其中 $c_i$ 为常数, $A_i \in \mathscr{F}$, $1 \leqslant i \leqslant n$, $\mathbf{1}_A$ 表集合 $A$ 上的**示性函数**, 即是 $\mathbf{1}_A(x) = 1$ 当 $x \in A$ 而 $\mathbf{1}_A(x) = 0$ 当 $x \overline{\in} A$, 则称 $f$ 为 $\mathscr{F}$ **简单函数**. 当不致混淆时, $\mathscr{F}$ 可测函数 (或 $\mathscr{F}$ 简单函数) 简称为**可测函数** (或简单函数). 易证: 任何非负 $\mathscr{F}$ 可测函数 $f$, 均可表为一串单增的 $\mathscr{F}$ 简单函数 $\{f_n\}$ 的极限: $f = \lim\limits_{n\to\infty} f_n$, 其中 $f_n \leqslant f_{n+1} (n \geqslant 1)$.

**定理 1.3 (单调类定理)** 设 $\Omega$ 是任一给定的集合, $\mathfrak{M}$ 是 $\Omega$ 上的一个 $\pi$ 系, $\mathscr{H}$ 是定义在 $\Omega$ 上的满足下列条件的实值函数构成的向量空间:

(1) 常数 $1 \in \mathscr{H}$, 且对任何 $A \in \mathfrak{M}, \mathbf{1}_A \in \mathscr{H}$;

(2) $\mathscr{H}$ 包含了一切含于 $\mathscr{H}$ 的非负的单增的函数列 $\{f_n\}$ 的实值极限函数 (相应地, 有界的极限函数) $f = \lim\limits_{n\to\infty} f_n$, 即

"$f_n \in \mathscr{H}$, $0 \leqslant f_n \leqslant f_{n+1}$, $(n \geqslant 1)$, $f = \lim\limits_{n\to\infty} f_n$ 是实值函数 (相应地, 有界实值函数) $\Rightarrow f \in \mathscr{H}$",

则 $\mathscr{H}$ 包含了一切定义在 $\Omega$ 上的实值的 (相应地, 有界实值的) $\sigma(\mathfrak{M})$ 可测的函数.

**证** 令 $\mathscr{D} = \{A \subset \Omega : \mathbf{1}_A \in \mathscr{H}\}$. 由定理的假设 (1) 可知 $\Omega \in \mathscr{D}, \mathfrak{M} \subset \mathscr{D}$. 而 $\mathscr{H}$ 是向量空间, 故当 $A_1, A_2 \in \mathscr{D}$ 且 $A_1 \subset A_2$ 时可推出: $\mathbf{1}_{A_2 - A_1} = \mathbf{1}_{A_2} - \mathbf{1}_{A_1} \in \mathscr{H}$, 亦即 $A_2 - A_1 \in \mathscr{D}$. 又由 (2) 可知: 对 $\mathscr{D}$ 中任一单增序列 $\{A_n\}$ 总有

$$\mathbf{1}_{\bigcup\limits_{n=1}^{\infty} A_n} = \sup_{n \geqslant 1} \mathbf{1}_{A_n} = \lim_{n\to\infty} \mathbf{1}_{A_n} \in \mathscr{H},$$

即 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{D}$. 总之, $\mathscr{D}$ 是 $\Omega$ 上的一个 $d$ 系. 故由定理 1.2 得知 $\mathscr{D} \supset d(\mathfrak{M}) = \sigma(\mathfrak{M})$.

任取一个 $\sigma(\mathfrak{M})$ 可测函数 $f$, 令 $f^+ = f \vee 0, f^- = (-f) \vee 0$ 分别为 $f$ 的正部与负部, 它们都是非负的 $\sigma(\mathfrak{M})$ 可测的实值函数. $f^+$ 总可表为非负单增的简单函数列

$$\left\{ f_n^+ = \sum_{i=1}^{k_n} a_i^{(n)} \mathbf{1}_{A_i^{(n)}}, n \geqslant 1 \right\}$$

$(a_i^{(n)} \in \mathbf{R}, A_i^{(n)} \in \sigma(\mathfrak{M}), A_i^{(n)} \bigcap A_j^{(n)} = \varnothing$, 当 $i \neq j)$ 的极限. 由 $f_n^+ \in \mathscr{H}$ 及条件 (2) 可知 $f^+ \in \mathscr{H}$. 仿之可证 $f^- \in \mathscr{H}$. 于是, 由 $\mathscr{H}$ 是向量空间可得 $f = (f^+ - f^-) \in \mathscr{H}$. 定理证毕.

**引理 1.1** 设 $\Omega_1, \Omega_2$ 为任意两个给定的集合, $\mathfrak{M}$ 是 $\Omega_2$ 上的一个集合系, $f : \Omega_1 \mapsto \Omega_2$ 是由 $\Omega_1$ 到 $\Omega_2$ 的一个映射, 则

$$f^{-1}(\sigma(\mathfrak{M})) = \sigma(f^{-1}(\mathfrak{M})).$$

证明甚易, 读者可作为习题证明之.

今后恒用 $f^{-1}(\mathfrak{M}) \overset{\text{def.}}{=\!=} \{f^{-1}(A) : A \in \mathfrak{M}\}$, 而 $f^{-1}(A)$ 是 $f$ 在 $A$ 上的逆像集.

**引理 1.2** 设 $\Omega$ 是任一给定的集合, $\Gamma$ 是任一指标集, $(E_i, \mathscr{E}_i)$ 是可测空间, $\mathscr{F}_i$ 是 $\mathscr{E}_i$ 中的集合系, $\sigma(\mathscr{F}_i) = \mathscr{E}_i, f_i : \Omega \mapsto E_i, (i \in \Gamma)$. 再令

$$\mathscr{F} = \left\{ B : B = \bigcap_{i \in I} f_i^{-1}(A_i), A_i \in \mathscr{F}_i, I \text{ 是 } \Gamma \text{ 中之有限子集} \right\},$$

$$\sigma(f_i, i \in \Gamma) = \sigma\left( \bigcup_{i \in \Gamma} f_i^{-1}(\mathscr{E}_i) \right),$$

则

(1) $\mathscr{F}_i$ 是 $\pi$ 系 $(\forall i \in \Gamma) \Rightarrow \mathscr{F}$ 是 $\pi$ 系;

(2) $\sigma(\mathscr{F}) = \sigma(f_i, i \in \Gamma)$.

证明甚易, 读者可作为习题证明之.

**定理 1.4 (单调类定理)** 设 $\Omega, \Gamma, I, \mathscr{F}, \mathscr{F}_i$ 和 $f_i$ 如引理 1.2 中所定义, 且 $\mathscr{F}_i$ 是 $\pi$ 系. 若 $\mathscr{H}$ 是定义在 $\Omega$ 上的一些实值函数所构成的向量空间且满足下列条件:

(1) $1 \in \mathscr{H}$;

(2) "$h_n \in \mathscr{H}, 0 \leqslant h_n \leqslant h_{n+1} (n \geqslant 1), h = \sup\limits_{n \geqslant 1} h_n$ 有限 (相应地, 有界) $\Rightarrow h \in \mathscr{H}$";

(3) $\mathscr{H}^* \stackrel{\text{def.}}{=\!=} \left\{ f = \prod\limits_{i \in I} \mathbf{1}_{A_i}(f_i) : A_i \in \mathscr{F}_i, i \in I, I \text{ 是 } \Gamma \text{ 的有限子集} \right\} \subset \mathscr{H}$,

则 $\mathscr{H}$ 包含了一切 $\sigma(f_i, i \in \Gamma)$ 可测的实值 (相应地, 有界) 函数.

**证** 注意

$$\prod_{i \in I} \mathbf{1}_{A_i}(f_i) = \mathbf{1}_{\{\bigcap\limits_{i \in I} f_i^{-1}(A_i)\}},$$

则可得:

$$\mathscr{H}^* = \{\mathbf{1}_A : A \in \mathscr{F}\}.$$

由 (3) 得 $\{\mathbf{1}_A : A \in \mathscr{F}\} \subset \mathscr{H}$. 再注意 $\mathscr{F}$ 是 $\pi$ 系, 则由定理 1.3 立即得到定理 1.4.

**附注 1.1** 我们有时把定理 1.1 和 1.2 称为集合形式的单调类定理, 而把定理 1.3 和 1.4 称为函数形式的单调类定理.

## §2 测度的基本概念及性质

在这一节中, 我们要研究抽象空间中的**预测度**、**外测度**和测度的基本概念和简单性质.

**定义 2.1** 设 $\Omega$ 是任一给定的集合, $\mathscr{F}$ 是 $\Omega$ 上的一个集合系, $\tau : \mathscr{F} \mapsto [0, \infty]$. 称 $\tau$ 是 $\mathscr{F}$ 上的一个预测度, 如果 $\varnothing \in \mathscr{F}$ 且 $\tau(\varnothing) = 0$. 称 $\mathscr{F}$ 上的预测度 $\tau$ 是 $\mathscr{F}$ 上的外测度, 如果 $\tau$ 还满足:

(1) 单增性, 即当 $A_1, A_2 \in \mathscr{F}, A_1 \subset A_2$ 时总有 $\tau(A_1) \leqslant \tau(A_2)$;

(2) 半可加性, 即当 $\{A_n\} \subset \mathscr{F}, \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$ 时总有 $\tau\left(\bigcup\limits_{n=1}^{\infty} A_n\right) \leqslant \sum\limits_{n=1}^{\infty} \tau(A_n)$.

若 $\mathscr{F}$ 是 $\Omega$ 的全体子集时, 则简称 $\mathscr{F}$ 上的预测度 (相应地, 外测度) 为预测度 (相应地, 外测度).

特别地, 若 $\mathscr{F}$ 上的外测度 $\tau$ 还满足

(3) 若 $\{A_n\}$ 是 $\mathscr{F}$ 中的两两不交的集合列, 且 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$, 总有 $\tau\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = \sum\limits_{n=1}^{\infty} \tau(A_n)$ (可数可加性), 则称 $\tau$ 是 $\mathscr{F}$ 上的**测度**.

称测度 $\tau$ 是**概率测度**, 若 $\tau(\Omega) = 1$.

**定理 2.1** 设 $\mathscr{F}$ 是 $\Omega$ 上的一个集合系, $\tau$ 是 $\mathscr{F}$ 上的预测度, 对任何 $A \subset \Omega$, 定义

$$\mu(A) = \inf_{\substack{\bigcup\limits_i c_i \supset A, \\ c_i \in \mathscr{F}}} \sum_i \tau(c_i), \tag{2.1}$$

则 $\mu$ 是 $\Omega$ 上的一个外测度, 有时称 $\mu$ 是预测度 $\tau$ 按模式 (I) 产生的外测度. (注: 若 (2.1) 式右方不存在 $c_i \in \mathscr{F}, \bigcup\limits_i c_i \supset A$, 则按惯例, 定义 $\inf \varnothing = \infty$.)

**证** 由外测度的定义立刻可验证 $\mu$ 是一个外测度.

**定义 2.2** 设 $\mu$ 是 $\Omega$ 上的一个外测度, 称 $\Omega$ 的子集 $A$ 是 $\mu$ 可测的, 如果对任何 $A_1 \subset A, A_2 \subset (\Omega - A)$, 总有

$$\mu(A_1 \bigcup A_2) = \mu(A_1) + \mu(A_2).$$

用 $\sigma(\mu)$ 表全体 $\mu$ 可测集.

**命题 2.1** 若 $E_1, E_2 \in \sigma(\mu), E_1 \bigcap E_2 = \varnothing, A \subset \Omega$, 则

$$\mu(A \bigcap (E_1 \bigcup E_2)) = \mu(A \bigcap E_1) + \mu(A \bigcap E_2).$$

**证** 由 $\mu$ 可测集之定义立得此命题.

**定理 2.2** 设 $\mu$ 是 $\Omega$ 上的一个外测度, 则 $\sigma(\mu)$ 是 $\Omega$ 上的一个 $\sigma$ 代数, 而且 $\mu$ 在 $\sigma(\mu)$ 上的局限 $\mu|_{\sigma(\mu)}$ 是一个测度.

**证** 为证此定理, 只需证明下列四类:

(a) 如果 $\mu(A) = 0$, 则 $A \in \sigma(\mu)$;

(b) 如果 $B \in \sigma(\mu)$, 则 $B^c = (\Omega - B) \in \sigma(\mu)$;

(c) 如果 $\{E_n\} \subset \sigma(\mu)$, 则 $\sum\limits_{n=1}^{\infty} E_n \in \sigma(\mu)$, $\bigcap\limits_{n=1}^{\infty} E_n \in \sigma(\mu)$;

(d) 如果 $\{E_n\} \subset \sigma(\mu)$, 且 $\{E_n\}$ 两两不交, 则

$$\mu\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n). \tag{2.2}$$

现证 (a). 设 $\mu(A)=0, A_1\subset A, A_2\subset A^c$, 则

$$\begin{aligned}\mu(A_2)&\leqslant \mu(A_1\bigcup A_2)\leqslant \mu(A_1)+\mu(A_2)\\&\leqslant \mu(A)+\mu(A_2)=\mu(A_2),\end{aligned}$$

所以 $\mu(A_1\bigcup A_2)=\mu(A_1)+\mu(A_2)$, 从而 $A\in\sigma(\mu)$.

再证 (b). 用对称性立刻可证 (b).

第三, 对两个集合 $E_1, E_2$ 之并的特例来证 (c). 设 $E_1, E_2\in\sigma(\mu)$, $A$ 和 $B$ 是两个具有有限 $\mu$ 测度的满足下列条件的集合:

$$A\subset (E_1\bigcup E_2),\quad B\subset (E_1\bigcup E_2)^c.$$

因为

$$A\bigcup B=(A\bigcap E_1)\bigcup[(A\bigcup B)\bigcap(\Omega-E_1)],$$

而且

$$(A\bigcap E_1)\subset E_1,\quad (A\bigcup B)\bigcap(\Omega-E_1)\subset\Omega-E_1,\quad E_1 \text{ 是 } \mu \text{ 可测集},$$

所以

$$\mu(A\bigcup B)=\mu(A\bigcap E_1)+\mu((A\bigcup B)\bigcap(\Omega-E_1)). \tag{2.3}$$

但是

$$(A\bigcup B)\bigcap(\Omega-E_1)=\{A\bigcap(\Omega-E_1)\}\bigcup B,$$

而且

$$A\bigcap(\Omega-E_1)\subset E_2,\quad B\subset\Omega-E_2,\quad E_2 \text{ 是 } \mu \text{ 可测集},$$

所以

$$\mu((A\bigcup B)\bigcap(\Omega-E_1))=\mu(A\bigcap(\Omega-E_1))+\mu(B). \tag{2.4}$$

再用 $E_1\in\sigma(\mu)$, 得

$$\mu(A\bigcup E_1)+\mu(A\bigcap(\Omega-E_1))=\mu(A). \tag{2.5}$$

由 (2.3)~(2.5) 得

$$\mu(A\bigcup B)=\mu(A)+\mu(B), \tag{2.6}$$

所以 $E_1\bigcup E_2$ 是 $\mu$ 可测集.

第四, 设 $\{E_n\}$ 是两两不交的 $\mu$ 可测集, 要证 (2.2) 成立 (即 (d) 成立), 而且 $\bigcup\limits_{n=1}^{\infty} E_n$ 和 $\bigcap\limits_{n=1}^{\infty} E_n$ 都是 $\mu$ 可测集. 记

$$E=\bigcup_{n=1}^{\infty} E_n,$$

且任取 $A\subset E$ 且 $B\subset\Omega-E$, 反复利用第三步, 可证: 对任何正整数 $n$, $\bigcup\limits_{i=1}^{n}E_i$ 都是 $\mu$ 可测集. 因此, 由 $B\subset\Omega-E\subset\Omega-\bigcup\limits_{i=1}^{n}E_i$, $A\bigcap\left(\bigcup\limits_{i=1}^{n}E_i\right)\subset\bigcup\limits_{i=1}^{n}E_i$ 得

$$\begin{aligned}\mu(A\bigcup B)&\geqslant\mu\left(\left[A\bigcap\left(\bigcup_{i=1}^{n}E_i\right)\right]\bigcup B\right)\\&=\mu\left(A\bigcap\left(\bigcup_{i=1}^{n}E_i\right)\right)+\mu(B).\end{aligned}\tag{2.7}$$

由于 $E_n,E_{n-1},\cdots,E_1$ 是两两不交的 $\mu$ 可测集, 所以由命题 2.1 得:

$$\begin{aligned}\mu\left(A\bigcap\left(\bigcup_{i=1}^{n}E_i\right)\right)&=\mu\left(\left[A\bigcap\left(\bigcup_{i=1}^{n-1}E_i\right)\right]\bigcup[A\bigcap E_n]\right)\\&=\mu\left(A\bigcap\left(\bigcup_{i=1}^{n-1}E_i\right)\right)+\mu(A\bigcap E_n)\\&=\mu\left(A\bigcap\left(\bigcup_{i=1}^{n-2}E_i\right)\right)+\mu(A\bigcap E_{n-1})+\mu(A\bigcap E_n)\\&=\cdots\\&=\sum_{i=1}^{n}\mu(A\bigcap E_i).\end{aligned}\tag{2.8}$$

将 (2.8) 代入 (2.7) 得知对任何正整数 $n$ 有:

$$\mu(A\bigcup B)\geqslant\sum_{i=1}^{n}\mu(A\bigcap E_i)+\mu(B),\tag{2.9}$$

所以

$$\mu(A\bigcup B)\geqslant\sum_{n=1}^{\infty}\mu(A\bigcap E_n)+\mu(B).\tag{2.10}$$

再用外测度 $\mu$ 的半可加性得:

$$\begin{aligned}\mu(A\bigcup B)&\geqslant\sum_{n=1}^{\infty}\mu(A\bigcap E_n)+\mu(B)\\&=\mu\left(A\bigcap\left(\bigcup_{n=1}^{\infty}E_n\right)\right)+\mu(B)\\&\geqslant\mu(A)+\mu(B)\\&\geqslant\mu(A\bigcup B).\end{aligned}\tag{2.11}$$

因此 (2.11) 中的所有的 $\geqslant$ 都应为 $=$. 由 $\mu(A\bigcup B)=\mu(A)+\mu(B)$ 知 $\bigcup\limits_{n=1}^{\infty}E_n$ 是 $\mu$ 可测集. 在 (2.11) 中取 $B=\varnothing, A=\bigcup\limits_{n=1}^{\infty}E_n$, 得

$$\mu\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\mu(E_n). \tag{2.12}$$

此即 (d) 成立. 由对称性知 $\bigcap\limits_{n=1}^{\infty}E_n=\left(\bigcup\limits_{n=1}^{\infty}E_n^{\mathrm{c}}\right)^{\mathrm{c}}$ 也是 $\mu$ 可测集.

最后, 对任意一串 $\mu$ 可测集 $\{E_n\}$, 要证 $\bigcup\limits_{n=1}^{\infty}E_n$ 和 $\bigcap\limits_{n=1}^{\infty}E_n$ 都是 $\mu$ 可测集. 事实上,

$$\bigcup_{n=1}^{\infty}E_n=\bigcup_{n=1}^{\infty}\left(E_n\bigcap\left(\Omega-\bigcup_{i<n}E_i\right)\right),$$

由前面三步的证明可知 $\{(E_n\bigcap(\Omega-\bigcup\limits_{i<n}E_i)), n\geqslant 1\}$ 是两两不交的 $\mu$ 可测集列, 再根据第四步, $\bigcup\limits_{n=1}^{\infty}E_n$ 是 $\mu$ 可测集. 利用对称性可知 $\bigcap\limits_{n=1}^{\infty}E_n=\left(\bigcup\limits_{n=1}^{\infty}E_n^{\mathrm{c}}\right)^{\mathrm{c}}$ 也是 $\mu$ 可测集. 定理证毕.

**定理 2.3** 设 $\mu$ 是 $\Omega$ 上的外测度, $T\subset\Omega, \{A_n\}\subset\sigma(\mu)$, 则

(1) $\{A_n\}$ 单增 $\Rightarrow \mu\left(T\bigcap\left(\bigcup\limits_{n=1}^{\infty}A_n\right)\right)=\sup\limits_{n\geqslant 1}\mu(T\bigcap A_n)$;

(2) $\{A_n\}$ 单降, 而且存在 $n_0$ 使 $\mu(A_{n_0}\bigcap T)<\infty \Rightarrow \mu\left(T\bigcap\left(\bigcap\limits_{n=1}^{\infty}A_n\right)\right)=\inf\limits_{n\geqslant 1}\mu(T\bigcap A_n)$.

**证** (1) 令 $E_n=A_n\bigcap\left(\Omega-\bigcup\limits_{i<n}A_i\right)$, 则 $\{E_n\}$ 是两两不交的 $\mu$ 可测集, 而且 $\bigcup\limits_{n=1}^{\infty}E_n=\bigcup\limits_{n=1}^{\infty}A_n$. 在 (2.11) (注意: (2.11) 式中一切 $\geqslant$ 号都是等号) 中取

$B = \varnothing, A = T\bigcap\left(\bigcup\limits_{n=1}^{\infty} E_n\right)$ 得

$$
\begin{aligned}
\mu\left(T\bigcap\left(\bigcup_{n=1}^{\infty} A_n\right)\right) &= \mu\left(T\bigcap\left(\bigcup_{n=1}^{\infty} E_n\right)\right)\\
&= \sum_{n=1}^{\infty} \mu\left(T\bigcap\left(\bigcup_{i=1}^{\infty} E_i\right)\bigcap E_n\right)\\
&= \sum_{n=1}^{\infty} \mu(T\bigcap E_n)\\
&= \sup_{n\geqslant 1} \sum_{i=1}^{n} \mu(T\bigcap E_i)\\
&= \sup_{n\geqslant 1} \mu\left(T\bigcap\left(\bigcup_{i=1}^{n} E_i\right)\right)\\
&= \sup_{n\geqslant 1} \mu(T\bigcap A_n).
\end{aligned}
\tag{2.13}
$$

(2) 由定理 2.2 知 $\sigma(\mu)$ 是一个 $\sigma$ 代数, 所以 $\bigcap\limits_{n=1}^{\infty} A_n$ 和 $A_m - \bigcap\limits_{n=1}^{\infty} A_n$ 都是 $\mu$ 可测集. 再用命题 2.1 和本定理的结论 (1) 可得:

$$
\begin{aligned}
&\mu\left(T\bigcap\left(\bigcap_{n=1}^{\infty} A_n\right)\right)\\
&= \mu(T\bigcap A_{n_0}) - \mu\left(T\bigcap\left(A_{n_0} - \bigcap_{n=1}^{\infty} A_n\right)\right)\\
&= \mu(T\bigcap A_{n_0}) - \mu\left(T\bigcap A_{n_0}\bigcap\left(\bigcup_{n=1}^{\infty} A_n^{\mathrm{c}}\right)\right)\\
&= \mu(T\bigcap A_{n_0}) - \sup_{n\geqslant 1} \mu(T\bigcap A_{n_0}\bigcap A_n^{\mathrm{c}})\\
&= \mu(T\bigcap A_{n_0}) - \sup_{n\geqslant 1} \mu(T\bigcap A_{n_0} - T\bigcap A_{n_0}\bigcap A_n)\\
&= \inf_{n\geqslant 1} \mu(T\bigcap A_{n_0}\bigcap A_n)\\
&= \inf_{n\geqslant 1} \mu(T\bigcap A_n).
\end{aligned}
$$

定理证毕.

**推论 2.1** 设 $\mathscr{F}$ 是 $\Omega$ 上的 $\sigma$ 代数, $\mu$ 是 $\mathscr{F}$ 上的测度, $\{A_n\} \subset \mathscr{F}$, 则

(1) 若 $\{A_n\}$ 单增, 则 $\mu\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} \mu(A_n)$;

(2) 若 $\{A_n\}$ 单降, 且存在 $n_0$ 使 $\mu(A_{n_0}) < \infty$, 则 $\mu\left(\bigcap\limits_{n=1}^{\infty} A_n\right) = \lim\limits_{n\to\infty} \mu(A_n)$.

**定理 2.4** 设 $\{\mu_n, n\in\Gamma\}$ 是一族外测度, $\Gamma$ 是任一指标集, 则

$$\mu \stackrel{\text{def.}}{=\!=} \sup_{n\in\Gamma}\mu_n$$

也是一个外测度.

**证** 由外测度之定义立即可证得此定理.

## §3 距离空间上的测度

设 $(\Omega,\rho)$ 是一个距离空间, $\rho$ 是 $\Omega$ 上的一个距离. 对任何 $x,y\in\Omega, A,B\subset\Omega$, $\rho(x,y)$ 表 $x$ 和 $y$ 之间的距离; $\rho(x,A)\stackrel{\text{def.}}{=\!=}\inf\limits_{y\in A}\rho(x,y)$ 表点 $x$ 列集合 $A$ 的距离; $\rho(A,B)\stackrel{\text{def.}}{=\!=}\inf\limits_{x\in A}\rho(x,B)$ 表集合 $A$ 与集合 $B$ 之间的距离. $\operatorname{diam}(A)=\sup\limits_{x,y\in A}\rho(x,y)$ 表集合 $A$ 的直径, 空集 $\varnothing$ 的直径 $\operatorname{diam}(\varnothing)$ 定义为 0.

**定义 3.1** 设 $(\Omega,\rho)$ 是一个**距离空间**, $\mathscr{F}$ 是 $\Omega$ 上的一个子集系, $\tau$ 是 $\mathscr{F}$ 上的一个预测度, 对任何 $B\subset\Omega, \varepsilon>0$, 令

$$\mu_\varepsilon(B)=\inf\left\{\sum_{i=1}^{\infty}\tau(c_i): c_i\in\mathscr{F}\bigcup_i c_i\supset B, \operatorname{diam}(c_i)\leqslant\varepsilon\right\},$$

易证 $\mu_\varepsilon$ 是 $\Omega$ 上的一个外测度. 再令

$$\mu(B)=\sup_{\varepsilon>0}\mu_\varepsilon(B)\quad(B\subset\Omega),$$

用定理 2.4 可证 $\mu$ 也是 $\Omega$ 上的一个外测度, 称之为由预测度 $\tau$ 按模式 (Ⅱ) 产生的外测度.

显然, 当 $\varepsilon$ 下降时 $\mu_\varepsilon$ 上升, 故

$$\mu(B)=\sup_{\varepsilon>0}\mu_\varepsilon(B)=\lim_{\varepsilon\to 0}\mu_\varepsilon(B)\quad(B\subset\Omega).$$

**定义 3.2** 设 $(\Omega,\rho)$ 是一个距离空间, $A,B\subset\Omega$, 若 $\rho(A,B)>0$, 则称 $A$ 和 $B$ 是一对**隔离集**. 称 $\Omega$ 上的外测度 $\mu$ 是 $\Omega$ 上的一个距离外测度, 如果对任何一对隔离集 $A$ 和 $B$, 都有 $\mu(A\bigcup B)=\mu(A)+\mu(B)$.

**定理 3.1** 距离空间 $(\Omega,\rho)$ 上的任何一个预测度 $\tau$ 按模式 (Ⅱ) 产生的外测度 $\mu$ 都是**距离外测度**.

**证** 为证定理, 只需证明对任何一对隔离集 $A$ 和 $B$, 都有

$$\mu(A\bigcup B)\geqslant\mu(A)+\mu(B). \tag{3.1}$$

不失普遍性, 可设 $\mu(A\bigcup B)<\infty$ 且 $\operatorname{diam}(\Omega)>0$, 否则 (3.1) 自然成立.

取 $\delta > 0, \delta_1 > 0, \delta_2 > 0$, 使

$$\rho(A,B) \geqslant \delta, \quad \mathrm{diam}(\Omega) \geqslant \delta_i > 0, (i=1,2). \text{ 令 } \eta = \min\left\{\delta_1, \delta_2, \frac{\delta}{2}\right\}.$$

由 $\mu(A \bigcup B) < \infty$ 及 $\mu$ 是预测度 $\tau$ 按模式 (II) 产生的外测度可知:

$$\mu(A \bigcup B) \geqslant \inf\left\{\sum_i \tau(c_i): \ \mathrm{diam}(c_i) \leqslant \eta, \bigcup_i c_i \supset (A \bigcup B).\right\}$$

因此, 对任何 $\varepsilon > 0$, 可取 $\{c_i\}$ 使

$$\begin{aligned} &\bigcup_i c_i \supset (A \bigcup B); \\ &\mathrm{diam}(c_i) \leqslant \eta \quad (i=1,2,\cdots); \\ &\sum_i \tau(c_i) \leqslant \mu(A \bigcup B) + \varepsilon. \end{aligned} \tag{3.2}$$

谬设存在一个 $i$, 使

$$c_i \bigcap A \neq \varnothing \neq c_i \bigcap B, \tag{3.3}$$

则可取 $a_0 \in c_i \bigcap A, b_0 \in c_i \bigcap B$, 使

$$\begin{aligned} \rho(a_0, b_0) &\leqslant \mathrm{diam}(c_i) \leqslant \eta \leqslant \frac{\delta}{2} \\ &\leqslant \frac{1}{2}\rho(A,B) \leqslant \frac{1}{2}\rho(a_0, b_0). \end{aligned} \tag{3.4}$$

因此 $\rho(a_0, b_0) = 0 = \eta$. 这与 $\eta$ 的取法矛盾, 所以, 没有一个 $i$ 会使 (3.3) 成立. 再令

$$A_i = \begin{cases} c_i, & \text{若 } c_i \bigcap A \neq \varnothing, \\ \varnothing, & \text{反之}; \end{cases}$$

$$B_i = \begin{cases} c_i, & \text{若 } c_i \bigcap B \neq \varnothing, \\ \varnothing, & \text{反之} \end{cases}$$

$$(i = 1, 2, \cdots),$$

则

$$\begin{aligned} &\mathrm{diam}(A_i) \leqslant \eta \leqslant \delta_1, \quad \mathrm{diam}(B_i) \leqslant \eta \leqslant \delta_2 \quad (i=1,2,\cdots), \\ &\bigcup_{i=1}^{\infty} A_i \supset A, \quad \bigcup_{i=1}^{\infty} B_i \supset B. \end{aligned} \tag{3.5}$$

由于没有一个 $i$ 会使 (3.3) 成立, 所以由 $A_i$ 和 $B_i$ 的定义得知: 对同一个 $i, A_i$ 和 $B_i$ 中至少有一个是空集, 所以

$$\begin{aligned}\tau(A_i)+\tau(B_i) &= \max\{\tau(A_i),\tau(B_i)\}\\ &\leqslant \max\{\tau(c_i),\tau(\varnothing)\}\\ &= \tau(c_i) \quad (i=1,2,\cdots).\end{aligned} \tag{3.6}$$

由 (3.5)、(3.6) 和 (3.2) 得:

$$\begin{aligned}\mu_{\delta_1}(A)+\mu_{\delta_2}(B) &\leqslant \sum_{i=1}^{\infty}\tau(A_i)+\sum_{i=1}^{\infty}\tau(B_i)\\ &\leqslant \sum_{i=1}^{\infty}\tau(c_i)\leqslant \mu(A\bigcup B)+\varepsilon.\end{aligned} \tag{3.7}$$

由 $\varepsilon>0, \delta_1>0$ 和 $\delta_2>0$ 可任意接近 0 可知

$$\mu(A)+\mu(B)\leqslant \mu(A\bigcup B).$$

定理证毕.

**定义 3.3** 设 $(\Omega,\rho)$ 是一个距离空间, $\mathscr{G}$ 是 $\Omega$ 中的全体开集所构成的集合系, 由 $\mathscr{G}$ 产生的 $\sigma$ 代数 $\sigma(\mathscr{G})$ 称为 $\Omega$ 上的 Borel $\sigma$ 代数, 记之为 $\mathscr{B}(\Omega)$, 其中每一集均称为 Borel 集.

**定理 3.2** 设 $\mu$ 是距离空间 $(\Omega,\rho)$ 上的任意一个距离外测度, 则每一 Borel 集皆 $\mu$ 可测, 即

$$\mathscr{B}(\Omega)\subset\sigma(\mu).$$

**证** 显然 $\varnothing,\Omega\in\sigma(\mu)$. 任取 $\Omega$ 中的一个不等于 $\Omega$ 的非空闭集 $F$, 若能证明 $F\in\sigma(\mu)$, 则定理得证. 事实上, 任意一个这样的闭集 $F$, 必存在两个非空的集合 $A$ 和 $B$, 满足

$$A\subset F, \quad B\subset\Omega-F.$$

下面我们要构造一列单增的集合 $\{B_n\}$ 使

(1) $B=\bigcup\limits_{n=1}^{\infty}B_n$;

(2) 对每个 $n\geqslant 1, B_n$ 与 $\Omega-B_{n+1}$ 是一对隔离集;

(3) 对每个 $n\geqslant 1, A$ 和 $B_n$ 也是一对隔离集.

令

$$B_n=\left\{x\in B:\rho(x,F)>\frac{1}{n}\right\} \quad (n\geqslant 1),$$

往证 $\{B_n\}$ 即为所求. 显然 $\{B_n\}$ 单增. 由于 $F$ 是闭集且 $B \subset \Omega - F$, 所以 (1) 成立. 由于 $F$ 是闭集, $A \subset F$, 所以由 $B_n$ 的定义得知: $B_n$ 与 $A$ 是一对隔离集 $(\forall n \geqslant 1)$, 此即 (3) 成立. 再证 (2) 成立. 任取 $b \in B_n$, $c \in \Omega - B_{n+1}, (n \geqslant 1)$. 由 $B_{n+1}$ 的定义有

$$\inf_{y \in F} \rho(c, y) \leqslant \frac{1}{n+1}.$$

所以可取一点 $f \in F$, 使

$$\rho(c, f) < 1 \Big/ \left(n + \frac{1}{2}\right).$$

谬设

$$\rho(b, c) \leqslant \frac{\frac{1}{2}}{n\left(n + \frac{1}{2}\right)},$$

则

$$\begin{aligned} \rho(b, F) \leqslant \rho(b, f) &\leqslant \rho(b, c) + \rho(c, f) \\ &\leqslant \frac{\frac{1}{2}}{n\left(n + \frac{1}{2}\right)} + \frac{1}{\left(n + \frac{1}{2}\right)} = \frac{1}{n}. \end{aligned}$$

这与 $b \in B_n$ 矛盾, 所以

$$\rho(b, c) > \frac{\frac{1}{2}}{n\left(n + \frac{1}{2}\right)}.$$

注意 $b$ 和 $c$ 分别是 $B_n$ 和 $\Omega - B_{n+1}$ 中任何一点, 所以

$$\rho(B_n, \Omega - B_n) \geqslant \frac{\frac{1}{2}}{n\left(n + \frac{1}{2}\right)} \quad (n \geqslant 1).$$

此即 (2) 成立.

利用 $\{B_n\}$ 的构造及其性质, 我们有:

$$\begin{aligned} \mu(A \bigcup B) &\geqslant \sup_{n \geqslant 1} \mu(A \bigcup B_n) \\ &= \sup_{n \geqslant 1} \{\mu(A) + \mu(B_n)\} \\ &= \mu(A) + \sup_{n \geqslant 1} \mu(B_n) \\ &= \mu(A) + \mu(B). \end{aligned}$$

因此 $\mu(A \bigcup B) = \mu(A) + \mu(B)$, 从而 $F \in \sigma(\mu)$. 定理证毕.

## §4 $N$ 维欧氏空间中的 L-S 测度

对 $N$ 维欧氏空间 $\mathbf{R}^N$, 其中的点, 用 $x=(x_1,\cdots,x_N), y=(y_1,\cdots,y_N), a=(a_1,\cdots,a_N)$ 和 $b=(b_1,\cdots,b_N)$ 表示. 记

$$(a,b]=\{x\in\mathbf{R}^N: a_i<x_i\leqslant b_i, 1\leqslant i\leqslant N\}$$

为 $\mathbf{R}^N$ 中的左开右闭 $N$ 维区间 ($N$ 维矩形)

$$\mathscr{I}_N=\{(a,b]: a,b\in\mathbf{R}^N, a\leqslant b\}$$

为 $\mathbf{R}^N$ 中全体左开右闭的 $N$ 维区间的全体, 此处 $a\leqslant b$, 意即 $a_i\leqslant b_i(i=1,\cdots,N)$.$a\geqslant b, a<b, a>b$ 的意义类似. $(a,b), [a,b], [a,b)$ 的定义仿 $(a,b]$.

**定义 4.1** 设 $F:\mathbf{R}^N\mapsto\mathbf{R}$. 称 $F$ 是 $(\text{L-S})_N$ 函数 ($N$ 维 Lebesgue-Stieltjes 函数), 如果

(1) 如果 $F(x_1,\cdots,x_N)$ 对每个变量皆右连续;

(2) 任取 $(a,b]\in\mathscr{I}_N$, 均有:

$$\begin{aligned}&\mu_F((a,b])\overset{\text{def.}}{=\!=}F(b_1,\cdots,b_n)-[F(a_1,b_2,\cdots,b_N)\\&+F(b_1,a_2,b_3,\cdots,b_N)+\cdots+F(b_1,\cdots,b_{N-1},a_N)]\\&+\cdots+(-1)^N F(a_1,\cdots,a_N)\geqslant 0.\end{aligned}\tag{4.1}$$

显然 $F(x_1,\cdots,x_N)=x_1\cdots x_N$ 是 $(\text{L-S})_N$ 函数, 这个特殊的 $(\text{L-S})_N$ 函数称为 $L_N$ 函数 ($N$ 维 Lebesgue 函数).

称 $(\text{L-S})_N$ 函数 $F$ 是**测度有界**的, 如果

$$\sup_{\substack{a,b\in\mathbf{R}^N\\a\leqslant b}}\mu_F((a,b])\leqslant M<\infty;\tag{4.2}$$

称非负的、测度有界的 $(\text{L-S})_N$ 函数 $F$ 是**标准的 $(\text{L-S})_N$ 函数**, 如果它还满足

$$\lim_{\substack{x_k\to\infty\\1\leqslant k\leqslant N}}F(x_1,\cdots,x_N)=1,\tag{4.3}$$

而且对任意一个 $1\leqslant k\leqslant N$, 有

$$\lim_{x_k\to-\infty}F(x_1,\cdots,x_N)=0.\tag{4.4}$$

**附注 4.1** $(\text{L-S})_1$ 函数简记为 (L-S) 函数.

**定理 4.1** 设 $F(x_1,\cdots,x_N)$ 是标准的 $(\text{L-S})_N$ 函数, 对任意的左开右闭的 $N$ 维区间 $(a,b]\in\mathscr{I}_N$, 定义 $\mu_F((a,b])$ 如 (4.1) (注意: 集合系 $\mathscr{I}_N$ 是**半环**), 则 $\mu_F$ 是半环 $\mathscr{I}_N$ 上的一个不大于 1 的测度, 从而它可以唯一地扩张到 $\sigma(\overline{\mathscr{I}}_N)=\mathscr{B}(\mathbf{R}^N)$ 上去而成一个概率测度, 这里 $\mathscr{B}(\mathbf{R})$ 表示 $\mathbf{R}^N$ 中一切 Borel 集所构成的 Borel $\sigma$ 代数.

**证** 为证此定理, 只需证明如上定义的 $\mu_F$ 是 $\mathscr{I}_N$ 上的一个有限测度. 显然, $\mu_F$ 在 $\mathscr{I}_N$ 上非负有限, 因此为证 $\mu_F$ 是 $\mathscr{I}_N$ 上的有限测度, 只需证明: $\mu_F$ 在 $\mathscr{I}_N$ 上有完全可加性. 为此, 先证明几个引理.

**引理 4.1** 设 $\{I_i,1\leqslant i\leqslant n\}$ 两两不交且含于 $\mathscr{I}_N$, 而且 $I=\bigcup\limits_{i=1}^{n}I_i\in\mathscr{I}_N$, 则

$$\mu_F(I)=\sum_{i=1}^{n}\mu_F(I_i). \tag{4.5}$$

**证** 用归纳法. 当 $n\leqslant 2$ 时, 由 $\mu_F$ 的定义直接验算可知 (4.5) 成立. 设 $n=k-1\geqslant 2$ 时 (4.5) 成立. 要证 $n=k$ 时 (4.5) 亦成立. 事实上, 令 $I=(a,b]$, 总存在 $1\leqslant m\leqslant k$, 使 $I_m=(c,b],c\geqslant a,c\neq a$. 不失普遍性可令 $m=1,c_1>a_1(c=(c_1,\cdots,c_N),a=(a_1,\cdots,a_N))$. 这时令

$$J=(c_1,b_1;a_2,b_2;\cdots;a_N,b_N],$$
$$K=(a_1,c_1;a_2,b_2;\cdots;a_N,b_N],$$

则 $I_1\subset J,J\bigcap K=\varnothing,I=J\bigcup K$, 且存在一个 $l>1$, 使 $I_l\subset K$, 不妨令 $l=2$. 所以

$$J=J\bigcap I=JI_1\bigcup JI_3\bigcup JI_4\bigcup\cdots\bigcup JI_k,$$
$$K=K\bigcap I=KI_2\bigcup KI_3\bigcup KI_4\bigcup\cdots\bigcup KI_k,$$

从而

$$\begin{aligned}\mu_F(I)&=\mu_F(J)+\mu_F(K)\\&=\mu_F(I_1)+\sum_{i=3}^{k}\mu_F(JI_i)\\&\quad+\mu_F(I_2)+\sum_{i=3}^{k}\mu_F(KI_i)\\&=\sum_{i=1}^{k}\mu_F(I_i).\end{aligned}$$

归纳法完成, 引理 4.1 证毕.

**引理 4.2**　设 $\{I, I_1, I_2, \cdots\} \subset \mathscr{I}_N$, $\{I_1, I_2, \cdots\}$ 两两不交且 $I \supset \bigcup\limits_{n=1}^{\infty} I_n$, 则

$$\mu_F(I) \geqslant \sum_{n=1}^{\infty} \mu_F(I_n). \tag{4.6}$$

**证**　显然, 对任意正整数 $k$, 都有

$$I \supset \bigcup_{n=1}^{k} I_n, \quad I - \bigcup_{n=1}^{k} I_n = \bigcap_{n=1}^{k} (I - I_n).$$

而 $\overline{\mathscr{I}}_N$ 是半环, 所以

$$I - I_n = \sum_{i=1}^{r_n} I_{n,i},$$
$$\{I_{n,i}, 1 \leqslant i \leqslant r_n\} \text{ 两两不交且 } \{I_{n,i}\} \subset \overline{\mathscr{I}}_N,$$

所以

$$I - \bigcup_{n=1}^{k} I_n = \bigcup_{j=1}^{m_k} J_j,$$
$$\{J_1, \cdots, J_{m_k}\} \text{ 两两不交且 } \{J_j\} \subset \overline{\mathscr{I}}_N,$$

故

$$\mu_F(I) = \sum_{n=1}^{k} \mu_F(I_n) + \sum_{j=1}^{m_k} \mu_F(J_j) \geqslant \sum_{n=1}^{k} \mu_F(I_n).$$

由 $k$ 的任意性可知引理 4.2 成立.

**引理 4.3**　若 $\{I, I_1, I_2, \cdots\} \subset \mathscr{I}_N$, $I \subset \bigcup\limits_{n=1}^{\infty} I_n$, 则 $\mu_F(I) \leqslant \sum\limits_{n=1}^{\infty} \mu_F(I_n)$.

**证**　令 $I = (a, b]$, $I_n = (a^{(n)}, b^{(n)}]$. 由于 $F(x_1, \cdots, x_N)$ 对每个自变量均右连续, 所以

$$\lim_{b_i \downarrow a_i} \mu_F(I) = 0 \quad (i = 1, \cdots, N), \tag{4.7}$$

从而对任意给定的 $\varepsilon > 0$, 都存在 $\overline{a}, \overline{b}^{(n)} \in \mathbf{R}^N, \overline{a} > 0, \overline{b}^{(n)} > 0$, 使

$$\mu_F((a + \overline{a}, b]) \geqslant \mu_F((a, b]) - \varepsilon, \tag{4.8}$$

$$\mu_F((a^{(n)}, b^{(n)} + \overline{b}^{(n)}]) \leqslant \mu_F((a^{(n)}, b^{(n)}]) + \frac{\varepsilon}{2^n}. \tag{4.9}$$

令

$$I' = (a + \overline{a}, b], \quad I'_n = (a^{(n)}, b^{(n)} + \overline{b}^{(n)}],$$

则

$$(\overline{I'}) \subset I \subset \bigcup_{n=1}^{\infty} I_n \subset \bigcup_{n=1}^{\infty} (I'_n)^0,$$

(此处及后文, $\overline{A}$ 表 $A$ 之闭包, 而 $A^0$ 表 $A$ 之**开核**.) 所以, 由有限覆盖定理得知: 存在一个有限的正整数 $k$, 使

$$(\overline{I'}) \subset \bigcup_{n=1}^{k} (I'_n)^0,$$

更有

$$I' \subset \bigcup_{n=1}^{k} I'_n.$$

所以

$$\begin{aligned} I' &= \bigcup_{n=1}^{k} I' I'_n \\ &= I' I'_1 \bigcup \left( \bigcup_{n=1}^{k-1} [(I' I'_1)^{\mathrm{c}} \cdots (I' I'_n)^{\mathrm{c}} (I' I'_{n+1})] \right) \end{aligned}$$

此处及后文中以 $A^{\mathrm{c}}$ 表示 $A$ 的补集. 由此推出:

$$\mu_F(I') \leqslant \mu_F(I'_1) + \cdots + \mu_F(I'_k). \tag{4.10}$$

所以由 (4.8)~(4.10) 得

$$\begin{aligned} \mu_F(I) &\leqslant \mu_F(I') + \varepsilon \\ &\leqslant \sum_{n=1}^{k} \left( \mu_F(I_n) + \frac{\varepsilon}{2^n} \right) + \varepsilon, \end{aligned}$$

而 $\varepsilon > 0$ 可以任意小, 所以

$$\mu_F(I) \leqslant \sum_{n=1}^{\infty} \mu_F(I_n).$$

引理 4.3 证毕.

由引理 4.1~4.3 得知 $\mu_F$ 是 $\mathscr{I}_N$ 上的有限测度. 定理 4.1 证毕.

**附注 4.2** 定理 4.1 中的由 $\mathscr{I}_N$ 唯一地扩张到 $\mathscr{B}(\mathbf{R}^N)$ 上去的测度 $\mu_F$, 仍用 $\mu_F$ 记之, 称之为 L-S 测度.

**附注 4.3** 对于积分

$$\int_{\mathbf{R}^N} f(x_1, \cdots, x_N) d\mu_F$$

有时记作

$$\int_{\mathbf{R}^N} f(x) d\mu_F, \text{ 或 } \int_{\mathbf{R}^N} f(x) \mu_F(dx), \text{ 或}$$
$$\int_{\mathbf{R}^N} f(x) dF, \text{ 或 } \int_{\mathbf{R}^N} f(x) F(dx),$$

此类积分, 通称为 (L – S) 积分.

## *§5 Hausdorff 测度

在这一节中, 简单地介绍一下 Hausdorff (外) 测度.

在本节中, 恒用 $\Phi$ 表示满足下述条件的函数 $\varphi$ 构成的函数族 (称 $\Phi$ 为测度函数族):

(1) $\varphi : (0, \delta) \mapsto (0, \infty), \delta$ 为一正数;

(2) $\varphi$ 是单增的、右连续的且 $\varphi(0+) = 0$. 此处及后文中恒用 $\varphi(x+0)(\varphi(x-0))$ 表 $\varphi$ 在 $x$ 的右 (左) 极限.

令 $\Phi_0$ 是 $\Phi$ 中的 "限制增长" 的子函数族:

$$\Phi_0 = \{\varphi \in \Phi : \text{ 存在 } K > 0, \text{ 使 } \varphi(2s)/\varphi(s) \leqslant K,$$
$$\text{对一切 } 0 < s < \frac{\delta}{2} \text{ 成立.}\}$$

**定义 5.1 (Hausdorff 外测度)** 任取 $\varphi \in \Phi, B \subset \mathbf{R}^N$, 定义

$$\varphi - m(B) = \liminf_{\varepsilon \downarrow 0} \left\{ \sum_{i=1}^{\infty} \varphi(\operatorname{diam}(G_i)) : G_i \in \mathscr{G}, \bigcup_{i=1}^{\infty} G_i \supset B, \operatorname{diam}(G_i) \leqslant \varepsilon \right\}, \tag{5.1}$$

此处 $\mathscr{G}$ 是 $\mathbf{R}^N$ 中全体开集, $\operatorname{diam}(G_i)$ 表 $G_i$ 的直径, 即是 $\operatorname{diam}(G_i) = \sup_{x,y \in G_i} |x - y|$, $|x-y|$ 表 $x$ 和 $y$ 在 $\mathbf{R}^N$ 中的欧氏距离, 特别地, 若 $N = 1$, $|x-y|$ 就是 $x$ 与 $y$ 之差的绝对值. (5.1) 式中极限符号下的 "$\varepsilon \downarrow 0$" 意指 $\varepsilon$ 单调下降从右方趋于零.

称 $\varphi - m(B)$ 为集合 $B$ 关于测度函数 $\varphi$ 的 Hausdorff 外测度. 特别地, 称 $s^\alpha - m(B)$ 为 $B$ 的 $\alpha$ 维 Hausdorff 外测度, 此处 $\alpha$ 是一个非负实数.

由定理 3.1 和定理 2.2 立得

**定理 5.1** Hausdorff 外测度 $\varphi - m(\bullet)$ 是一个距离外测度, 而 $\varphi - m(\bullet)$ 在 $\sigma(\varphi - m(\bullet))$ 上的局限是一个测度, 称之为 Hausdorff 测度.

**附注 5.1** 在一般的文献中, 多半只讨论 Hausdorff 外测度, 而不提 Hausdorff 测度, 并把 Hausdorff 外测度简称为 Hausdorff 测度.

**附注 5.2** 定理 5.1 中的 $\sigma(\varphi - m(\bullet))$ 包含了 $\mathbf{R}^N$ 中的全部 Borel 集 (请参见 [36] 第一章定理 2.6).

**附注 5.3** Hausdorff 外测度有许多等价性的定义, 有兴趣的读者请参看文献 [36] 定理 2.7.

下面介绍 Hausdorff 维数的定义及其简单性质.

**定义 5.2** 任取 $B \subset \mathbf{R}^N$, 定义

$$\dim(B) \overset{\text{def.}}{=\!=} \inf\{\alpha > 0 : s^\alpha - m(B) = 0\}$$

为 $B$ 的 **Hausdorff 维数**.

**定理 5.2** 对任何 $B \subset \mathbf{R}^N$, 总有:

(1) $\beta > \alpha, s^\alpha - m(B) < \infty \Rightarrow s^\beta - m(B) < \infty$;

(2) $\alpha > \dim(B) \Rightarrow s^\alpha - m(B) = 0$;

(3) $\alpha < \dim(B) \Rightarrow s^\alpha - m(B) = \infty$;

(4) $\alpha = \dim(B) \Rightarrow 0 \leqslant s^\alpha - m(B) \leqslant \infty$;

(5) $$\begin{aligned}\dim(B) &= \inf\{\alpha > 0 : s^\alpha - m(B) = 0\} \\ &= \inf\{\alpha > 0 : s^\alpha - m(B) < \infty\} \\ &= \sup\{\alpha > 0 : s^\alpha - m(B) = \infty\} \\ &= \sup\{\alpha > 0 : s^\alpha - m(B) > 0\}.\end{aligned}$$

**证** 由 Hausdorff 外测度和 Hausdorff 维数的定义即可验证此定理成立.

**定理 5.3 (Hausdorff 维数的 $\sigma$ 稳定性)** 对任何 $B_n \subset \mathbf{R}^N (n = 1, 2, \cdots)$, 总有

$$\dim\left(\bigcup_{n=1}^{\infty} B_n\right) = \sup_{n \geqslant 1} \dim(B_n).$$

**证** 显然, 由维数之定义有

$$\dim\left(\bigcup_{n=1}^{\infty} B_n\right) \geqslant \sup_{n \geqslant 1} \dim(B_n).$$

下面证明

$$\dim\left(\bigcup_{n=1}^{\infty} B_n\right) \leqslant \sup_{n \geqslant 1} \dim(B_n). \tag{5.2}$$

不失普遍性可设

$$\sup_{n\geqslant 1}\dim(B_n)<\infty.$$

令

$$\alpha_n=\dim(B_n),\quad \alpha=\sup_{n\geqslant 1}\alpha_n,$$

对任何 $\varepsilon>0$, 由定理 5.2 知

$$s^{\alpha+\varepsilon}-m(B_n)=0,\quad n=1,2,\cdots,$$

从而

$$s^{\alpha+\varepsilon}-m\left(\bigcup_{n=1}^{\infty}B_n\right)\leqslant\sum_{n=1}^{\infty}(s^{\alpha+\varepsilon}-m(B_n))=0.$$

于是

$$\dim\left(\bigcup_{n=1}^{\infty}B_n\right)\leqslant\alpha+\varepsilon.$$

由 $\varepsilon>0$ 可以任意小得知 (5.2) 成立. 定理证毕.

**定理 5.4** 对 $\mathbf{R}^N$ 中任何一个可数集 $B$, 总有 $\dim(B)=0$.

**证** 任取单点集 $\{x\}\subset\mathbf{R}^N$, 总有 $\dim(\{x\})=0$. 再用定理 5.3 立得定理 5.4.

## §6 习题及应用

1. 设 $(\Omega,\rho)$ 是距离空间, $A,B\subset\Omega,f:\Omega\to\mathbf{R}$. 称 $f$ 是 $A$ 与 $B$ 的一个隔离函数, 如果 $f(x)=0(\forall x\in A),f(x)=1(\forall x\in B)$ 且 $0\leqslant f\leqslant 1$.

(i) 若 $\rho(A,B)>0$, 则 $A$ 与 $B$ 有一致连续的隔离函数;

(ii) 若 $\rho(A,B)=0$, 但 $\overline{A}\bigcap\overline{B}=\varnothing$, 则 $A$ 与 $B$ 有连续的隔离函数;

(iii) 若 $\overline{A}\bigcap\overline{B}\neq\varnothing$, 则 $A$ 与 $B$ 没有连续的隔离函数;

(iv) 若 $A\bigcap B\neq\varnothing$, 则 $A$ 与 $B$ 没有隔离函数.

2. 设 $\mu$ 是 $\mathscr{B}(\mathbf{R}^2)$ 上的一个测度, 满足 $\mu(I)<\infty(\forall I\in\mathscr{I}_2)$. 试构造函数 $F(x_1,x_2)$, 它所产生的 $(\text{L-S})_2$ 测度就是 $\mu$.

3. (Hausdorff (外) 测度的等价定义. 设 $\varphi\in\Phi$ 是一个测度函数, $(\Omega,\rho)$ 是距

离空间. $\mathscr{G}(\mathscr{F})$ 是全体开 (闭) 集. 对任何 $B \subset \Omega$, 任何 $\varepsilon > 0$, 定义

$$\mu_\varepsilon^\varphi(B) = \inf\left\{\sum_{i=1}^\infty \varphi(\mathrm{diam}(G_i)) : G_i \in \mathscr{G}, \mathrm{diam}(G_i) \leqslant \varepsilon, \bigcup_{i=1}^\infty G_i \supset B\right\};$$

$$\gamma_\varepsilon^\varphi(B) = \inf\left\{\sum_{i=1}^\infty \varphi(\mathrm{diam}(F_i)) : F_i \in \mathscr{F}, \mathrm{diam}(F_i) \leqslant \varepsilon, \bigcup_{i=1}^\infty F_i \supset B\right\};$$

$$\sigma_\varepsilon^\varphi(B) = \inf\left\{\sum_{i=1}^\infty \varphi(\mathrm{diam}(B_i)) : B_i \subset \Omega, \mathrm{diam}(B_i) \leqslant \varepsilon, \bigcup_{i=1}^\infty B_i \supset B\right\};$$

$$\tau_\varepsilon^\varphi(B) = \inf\left\{\sum_{i=1}^\infty \varphi(\mathrm{diam}(B_i)) : B_i \subset \Omega, \mathrm{diam}(B_i) \leqslant \varepsilon, \bigcup_{i=1}^\infty B_i = B\right\}.$$

试证: 对任何 $\beta > \varepsilon$ 有

$$\mu_\beta^\varphi(B) \leqslant \gamma_\varepsilon^\varphi(B) = \sigma_\varepsilon^\varphi(B) = \tau_\varepsilon^\varphi(B) \leqslant \mu_\varepsilon^\varphi(B),$$

从而

$$\begin{aligned}\varphi - m(B) &= \lim_{\varepsilon\downarrow 0} \mu_\varepsilon^\varphi(B) = \lim_{\varepsilon\downarrow 0} \gamma_\varepsilon^\varphi(B) \\ &= \lim_{\varepsilon\downarrow 0} \sigma_\varepsilon^\varphi(B) = \lim_{\varepsilon\downarrow 0} \tau_\varepsilon^\varphi(B).\end{aligned}$$

4. 取 $(\Omega,\rho)$ 为特殊的距离空间 $(\mathbf{R}^N, \rho_N)$, 其中 $\rho_N$ 是 $\mathbf{R}^N$ 中的欧氏距离. 对任何 $B \subset \mathbf{R}^N, x \in \mathbf{R}^N, \lambda > 0$, 试证:

(i) $s^\alpha - m(x+B) = s^\alpha - m(B), \mathrm{dim}(x+B) = \mathrm{dim}(B)$;

(ii) $s^\alpha - m(\lambda B) = \lambda^\alpha(s^\alpha - m(B))$;

(iii) $\mathrm{dim}(\lambda B) = \mathrm{dim}(B)$;

(iv) $0 \leqslant \mathrm{dim}(B) \leqslant N$, $\mathrm{dim}(\mathbf{R}^N) = N$;

(v) $B$ 有内点 $\Rightarrow \mathrm{dim}(B) = N$.

5. 您知道衡量集合内涵丰富到何种程度有哪些标准? 并说明这些标准的优劣.

6. 设 $(\Omega,\rho)$ 是完备可分距离空间 (通常简称为 Polish 空间). $f:\Omega \mapsto \Omega$. 称

$$\mathrm{Lip}(f) \overset{\text{def.}}{=\!=} \sup_{\substack{x,y\in\Omega\\ x\neq y}} \frac{\rho(f(x),f(y))}{\rho(x,y)} \tag{6.1}$$

为 $f$ 的 Lipschitz 系数. 若 $\mathrm{Lip}(f) < 1$, 则称 $f$ 是压缩算子, 记为 $f \in \mathrm{Con}(\Omega)$. 设 $N \geqslant 2, K \subset \Omega$, 称 $K$ 为自相似集, 如果存在 $\{f_0, f_1, \cdots, f_{N-1}\} \subset \mathrm{Con}(\Omega)$, 使

$$K = \bigcup_{i=0}^{N-1} f_i(K), \tag{6.2}$$

$f_i(K)$ 表 $f_i$ 在 $K$ 上之像集.

称压缩算子 $f$ 是相似算子, 如果存在正实数 $a \in (0,1)$ 使

$$\rho(f(x), f(y)) = a\rho(x, y) \quad (\forall x, y \in \Omega).$$

记全体相似压缩算子为 $\mathrm{Sicon}(\Omega)$.

有时称满足 (6.2) 的集合 $K$ 为 $(f_0, f_1, \cdots, f_{N-1})$ 自相似集.

当 $\Omega \subset \mathbf{R}^d, \Omega$ 为紧集, 且 $\Omega = \overline{(\Omega^0)}, \{f_0, f_1, \cdots, f_{N-1}\} \subset \mathrm{Sicon}(\Omega), f_i(\Omega^0) \subset \Omega^0, f_i(\Omega^0) \bigcap f_j(\Omega^0) = \varnothing (i \neq j)$ 时, 任何 $\{f_0, f_1, \cdots, f_{N-1}\}$ 自相似集 $K$ 的 Hausdorff 维数为

$$\dim(K) = \alpha,$$

其中 $\alpha$ 是方程式

$$\sum_{i=0}^{N-1} \mathrm{Lip}(f_i)^\beta = 1, \quad 0 \leqslant \beta \leqslant d$$

的唯一解 (参见文献 [40].)

试证 $[0,1]$ 上的 Cantor 三分集 $C$ 是自相似集, 而且 $\dim(C) = \ln 2/\ln 3$.

7. Lebesgue 测度与 Hausdorff (外) 测度有何关系?

# 第二章　从实值随机变量到取值于 Banach 空间的随机元

## §1　随机变量及其分布, 母函数

**定义 1.1**　设 $(\Omega, \mathscr{F}, P)$ 是一个概率空间, 即是, $\Omega$ 是一个抽象集合, 称之为样本空间, $\mathscr{F}$ 是 $\Omega$ 上的 $\sigma$ 代数, $P$ 是 $\mathscr{F}$ 上的一个概率测度. 再设 $\mathbf{R}$ 是实数集, $\mathscr{B}(\mathbf{R})$ 是 $\mathbf{R}$ 上的 Borel $\sigma$ 代数, 即是由 $\mathbf{R}$ 中的全体开集所产生的 $\sigma$ 代数. $X$ 是定义在 $\Omega$ 上的实值函数, 如果对任何 $A \in \mathscr{B}(\mathbf{R})$, 都有

$$\{\omega \in \Omega : X(\omega) \in A\} \in \mathscr{F},$$

即 $X$ 是关于 $\sigma$ 代数 $\mathscr{F}$ 和 $\mathscr{B}(\mathbf{R})$ 的可测函数, 记为 $X \in \mathscr{F}/\mathscr{B}(\mathbf{R})$, 则称 $X$ 是 $(\Omega, \mathscr{F}, P)$ 上的实值**随机变量**. 在无混淆的情况下, 有时简称 $X$ 为随机变量.

显然, 对上述随机变量 $X$, 对任何实数 $x$, $\{\omega \in \Omega : X(\omega) \leqslant x\} \in \mathscr{F}$, $\{\omega \in \Omega : X(\omega) < x\} \in \mathscr{F}$, 即它们都是 $\mathscr{F}$ 可测集.

$\{\omega \in \Omega : X(\omega) \in A\}$ 是 $X$ 在 $A$ 上的逆像集, 简记之为 $X^{-1}(A)$.

$\{X \leqslant x\}$ 和 $X^{-1}((-\infty, x])$, $\{X < x\}$ 和 $X^{-1}((-\infty, x))$ 亦有类似的意义.

对上述随机变量 $X$ 而言, 称 $F(x) \stackrel{\text{def.}}{=\!=} P(X \leqslant x)(x \in \mathbf{R})$ 为 $X$ 的**分布函数**, 称 $P(X^{-1}(A))(A \in \mathscr{B}(\mathbf{R}))$ 为 $X$ 的分布, 记 $P(X^{-1}(A)) = (P \circ X^{-1})(A)$. 严格地说, 上述 $F(x)$ 和 $(P \circ X^{-1})(A)$ 分别为 $X$ 的分布函数在点 $x$ 之值和 $X$ 的分布在集 $A$ 上的测度值. 正像初等微积中那样, 在情况简单时, $f(x)$ 既表函数 $f$ 在点 $x$ 之函数值, 又表函数 $f$. 在情况复杂时, 随机变量 $X$ 的分布函数用 $F$ 或

$F(\bullet)$ 表示, 分布用 $P\circ X^{-1}$ 或 $(P\circ X^{-1})(\bullet)$ 表示. 如无混淆, $F(x)$ 既表分布函数, 又表它在 $x$ 之值.

注意: 随机变量 $X$ 的分布 $P\circ X^{-1}$ 是 $\mathscr{B}(\mathbf{R})$ 上的一个概率测度.

**定义 1.2** 设 $X$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的取值于 $\{x_0,x_1,\cdots\}\subset\mathbf{R}$ 的随机变量, 即其分布 $P\circ X^{-1}$ 仅在 $\{x_k\}$ 上可能有正测度, $(P\circ X^{-1})(\{x_k\})\geqslant 0,(k=0,1,\cdots)$, $\sum\limits_{k=0}^{\infty}P\circ X^{-1}(\{x_k\})=1$, 则称 $X$ 为离散的随机变量, 称一列非负的其和为 1 的实数 $\{p_0,p_1,p_2,\cdots\}$ 为一个 (离散的) 概率分布.

显然对一个取值于 $\{x_0,x_1,x_2,\cdots\}\subset\mathbf{R}$ 的离散的随机变量 $X$ 而言, 其分布 $P\circ X^{-1}$ 由概率分布 $\{p_0=P\circ X^{-1}(\{x_0\}),p_1=P\circ X^{-1}(\{x_1\}),\cdots\}$ 所决定, 其分布函数 $F$ 亦由概率分布 $\{p_0,p_1,\cdots\}$ 按下列公式所决定:

$$F(x)=\sum_{\{k:x_k\leqslant x\}}p_k,(x\in\mathbf{R}). \tag{1.1}$$

对于一般的实值随机变量 $X$ 而言, 其分布 $P\circ X^{-1}$ 或其分布函数 $F$ 决定了它的各种概率性质 (注意: 分布 $P\circ X^{-1}$ 与分布函数 $F$ 是相互唯一决定的), 而对取值于 $\{x_0,x_1,x_2,\cdots\}\subset\mathbf{R}$ 的离散随机变量 $X$ 而言, 其概率分布 $\{p_k=P\circ X^{-1}(\{x_k\}),k=0,1,2,\cdots\}$ 决定了它的各种概率性质.

由 Taylor 级数的理论的启示, 对于概率分布 $\{p_k,k=0,1,2,\cdots\}$ 而言, 以它为系数的 Taylor 级数, 在研究中起到重要作用, 因此我们引进

**定义 1.3** 设取值于 $\{x_0,x_1,\cdots\}\subset\mathbf{R}$ 的离散随机变量 $X$ 的概率分布为 $\{p_k,k=0,1,\cdots\}$, 则称

$$G(u)=\sum_{k=0}^{\infty}p_ku^k,\quad |u|\leqslant 1 \tag{1.2}$$

为 $X$ (或其概率分布 $\{p_k,k=0,1,\cdots\}$) 的母函数 (generating function).

显然 (1.2) 式右方之级数在 $|u|\leqslant 1$ 内绝对收敛, 在 $|u|<1$ 内解析.

离散随机变量 $X$ 的母函数 $G$ 比 $X$ 的分布函数 $F$, 有时使用起来更方便.

**例 1.1** 若离散随机变量 $X$ 的概率分布是二项分布: $\{p_k=b(k;n,p)\overset{\text{def.}}{=\!=}\binom{n}{k}p^kq^{n-l},k=0,1,\cdots,n\}$, 其中 $n$ 是固定的正整数, $p$ 是属于 $(0,1)$ 的固定的实数, $q=1-p$, $\binom{n}{k}$ 是 $n$ 个元素中取 $k$ 个的组合数, 则 $X$ 的母函数

$$G(u)=(q+pu)^n. \tag{1.3}$$

**例 1.2** 若离散随机变量 $X$ 的概率分布是参数为 $\lambda$ 的 Poisson 分布 $\{p_k = e^{-\lambda}\frac{\lambda^k}{k!}, k=0,1,\cdots\}$, 则 $X$ 的母函数为

$$G(u) = e^{-\lambda(1-u)}. \tag{1.4}$$

**定义 1.4** *若实值随机变量 $X$ 的分布函数 $F$ 绝对连续, 即存在一个 Lebesgue 可测函数 $f$, 使得对任何 $x \in \mathbf{R}$, 都有*

$$F(x) = \int_{-\infty}^{x} f(t)dt \tag{1.5}$$

*一般地, 若仍用 $P \circ X^{-1}$ 表 $X$ 的分布, 则对 $A \in \mathscr{B}(\mathbf{R})$, 有*

$$(P \circ X^{-1})(A) = \int_A f(t)dt. \tag{1.6}$$

*则我们称 $f$ 为 $X$ (或其分布函数 $F$) 的密度函数. 这时, 称 $X$ 是连续型的随机变量.*

注意: $X$ 是连续型的随机变量, 丝毫不意味着 $X$ 是连续函数, 因为 $X$ 的定义域 —— 样本空间 $\Omega$ 没有任何拓扑结构, 根本谈不上任何连续性, 连续型随机变量的意思是它的分布函数绝对连续.

连续型随机变量的最重要的代表, 是服从正态分布的随机变量, 它有密度函数

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-a}{\sigma}\right)^2}, \tag{1.7}$$

其中 $\sigma$ 为正实数, $a$ 为实数.

## §2 随机变量的独立性与测度的卷积

随机变量是用数值来刻画随机试验的结果, 有些随机试验的结果, 不一定用一个随机变量就能刻画. 例如, 从北京市人口中任意抽取一人, 测量其身高 (单位是 cm) $X$ 和体重 (单位是 kg) $Y$. 直观上看, $X$ 和 $Y$ 是随机变量, 但 $X$ 和 $Y$ 有某种统计相依性. 再从河南省和河北省的人口中, 各抽一人, 测量其身高 $X_1$ 和 $X_2$. 直观上看, $X_1$ 和 $X_2$ 是随机变量, 且 $X_1$ 与 $X_2$ 一般来说是统计独立的. 以后我们将会发现: 对于有某种统计相依性的随机变量 $X$ 和 $Y$, 只知道 $X$ 和 $Y$ 的各自的概率性质, 并不能完全刻画 $(X, Y)$ 作为一个整体的概率性质, 但是, 若 $X_1$ 和 $X_2$ 是统计独立的, 知道 $X_1$ 和 $X_2$ 各自的概率性质以后, 则 $(X_1, X_2)$ 作为一个整体, 其概率性质也就知道了. 因此, 研究随机变量之间的独立性, 就很有必要. 有了独立性, 对问题的研究, 可以大为简化. 如非事先声明, 言及随机变量, 必是概率空间 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量.

**定义 2.1** 设随机变量 $X_k$ 的分布函数为 $F_k, k=1,2,\cdots,N$. 记

$$F(x_1,\cdots,x_N)=P(X_1\leqslant x_1,\cdots,X_N\leqslant x_N), \tag{2.1}$$

称 $F$ 为 $(X_1,\cdots,X_N)$ 的联合分布函数. 如果

$$F(x_1,\cdots,x_N)=\prod_{k=1}^{N}F_k(x_k) \tag{2.2}$$

对一切 $\{x_1,\cdots,x_N\}\subset\mathbf{R}$ 都成立, 则称 $X_1,\cdots,X_N$ 是相互独立的.

(2.2) 的意思是 $X_1,\cdots,X_N$ 的各自的 (**边缘**) **分布函数** $F_k$ 之积等于 $(X_1,\cdots,X_N)$ 的**联合分布函数**.

**命题 2.1** 设 $X_1,\cdots,X_N$ 是 $N$ 个随机变量, 则下列陈述等价:

(1) $X_1,\cdots,X_N$ 是相互独立的;

(2) 对任何正整数 $k\leqslant N$, 任何正整数 $1\leqslant t_1<t_2<\cdots<t_k\leqslant N$, 任何 $\{x_1,\cdots,x_k\}\subset\mathbf{R}$, 都有

$$P\left(X_{t_1}\leqslant x_1,\cdots,X_{t_k}\leqslant x_k\right)=\prod_{i=1}^{k}F_{t_i}(x_i); \tag{2.3}$$

(3) 对任何正整数 $k\leqslant N$, 任何正整数 $1\leqslant t_1<t_2<\cdots<t_k\leqslant N$, 任何 $\{A_1,\cdots,A_k\}\subset\mathscr{B}(\mathbf{R})$, 都有

$$P\left(\bigcap_{i=1}^{k}X_{t_i}^{-1}(A_i)\right)=\prod_{i=1}^{k}P\left(X_{t_i}^{-1}(A_i)\right)=\prod_{i=1}^{k}\left(P\circ X_{t_i}^{-1}\right)(A_i). \tag{2.4}$$

注意: (2) 意即 $\{X_1,\cdots,X_N\}$ 中任取 $k\leqslant N$ 个随机变量 $\{X_{t_1},\cdots,X_{t_k}\}$, 它们都是相互独立的. (3) 意即 $\{X_{t_1},\cdots,X_{t_k}\}$ 的联合分布等于 $X_{t_i}(i=1,\cdots,k)$ 各自的分布之积.

**证** (1) $\Rightarrow$ (2). 设 (1) 成立. 令 $B=\{1,\cdots,N\}-\{t_1,\cdots,t_k\}$, 则

$$\begin{aligned}
&P\left(X_{t_1}\leqslant x_1,\cdots,X_{t_k}\leqslant x_k\right)\\
=&\lim_{\substack{y_s\to\infty\\ s\in B}}P\left(\bigcap_{i=1}^{k}\{X_{t_i}\leqslant x_i\},\bigcap_{s\in B}\{X_s\leqslant y_s\}\right)\\
=&\lim_{\substack{y_s\to\infty\\ s\in B}}\left(\prod_{i=1}^{k}F_{t_i}(x_i)\right)\left(\prod_{s\in B}F_s(y_s)\right)\\
=&\prod_{i=1}^{k}F_{t_i}(x_i),
\end{aligned}$$

此即 (2) 成立.

(2) ⇒ (3). 设 (2) 成立, $B$ 定义如前, $F^{\mathrm{T}}$ 是 $\{X_{t_1},\cdots,X_{t_k}\}$ 的联合分布函数所产生的 L-S 测度, $F_i$ 是由 $X_i$ 的分布函数所产生的 L-S 测度, $i=1,\cdots,n$, 则由 (2) 及单调系定理易证:

$$F^{\mathrm{T}}=F_{t_1}\times F_{t_2}\times\cdots\times F_{t_k}. \tag{2.5}$$

所以对任何 $\{A_1,\cdots,A_k\}\subset\mathscr{B}(\mathbf{R})$, 均有

$$\begin{aligned}&P\left(\bigcap_{i=1}^{k}X_{t_i}^{-1}(A_i)\right)\\&=\int_{A_1\times\cdots\times A_k}dF^{\mathrm{T}}=\int_{A_1\times\cdots\times A_k}d(F_{t_1}\times\cdots\times F_{t_k})\\&=\prod_{i=1}^{k}F_{t_i}(A_i)=\prod_{i=1}^{k}P\left(X_{t_i}^{-1}(A_i)\right).\end{aligned}$$

"(3) ⇒ (1)" 显然成立. 命题 2.1 证毕.

设 $E$ 是一个线性空间, $\mathscr{E}$ 是 $E$ 上的一个 $\sigma$ 代数, 于是得一个可测线性空间 $(E,\mathscr{E})$.

**定义 2.2** 设 $\mu_1,\mu_2,\cdots,\mu_N$ 均为 $(E,\mathscr{E})$ 上的有限测度, $\mu=\underset{i=1}{\overset{N}{X}}\mu_i$ 是它们的乘积测度. $\mu_1,\cdots,\mu_N$ 的卷积 (convolution) 定义为:

$$\begin{aligned}&(\mu_1*\cdots*\mu_N)(A)\\&=\mu\left(\{(x_1,\cdots,x_N)\in E:\sum_{i=1}^{N}x_i\in A\}\right)(A\in\mathscr{E}).\end{aligned} \tag{2.6}$$

有时我们记

$$\mu_1*\cdots*\mu_N=\underset{i=1}{\overset{N}{*}}\mu_i, \tag{2.7}$$

而 $\mu_1^{k*}$ 表示 $\mu_1$ 的 $k$ 重**卷积** $(k\geqslant 1)$. $(\mu_1^{k*})(A)$ 有时记为 $\mu_1(A)^{k*}$.

注意: $N$ 个有限测度的卷积仍为有限测度.

**命题 2.2** 设 $\mu_1,\cdots,\mu_N$ 都是 $(E,\mathscr{E})$ 上的有限测度.

(1) 对任何 $\mathscr{E}$ 可测的有界实值函数 $f$, 总有:

$$\begin{aligned}&\int_E(\mu_1*\cdots*\mu_N)(dx)f(x)\\&=\int_E\mu_1(dx_1)\cdots\int_E\mu_N(dx_N)f(x_1+\cdots+x_N).\end{aligned} \tag{2.8}$$

(2) 设 $\nu_i(x,A)$ 对任何固定的 $x\in E$ 作为 $A$ 的函数是 $\mathscr{E}$ 上的有限测度, 而对任何固定的 $A\in\mathscr{E}$ 作为 $x$ 的函数是 $\mathscr{E}$ 可测的, 则对任何 $A\in\mathscr{E}$, 总有

$$\mathop{*}_{i=1}^{N}\left(\int_E \mu_i(dx_i)\nu_i(x_i,A)\right)$$
$$=\int_E \mu_1(dx_1)\cdots\int_E \mu_N(dx_N)\left[\mathop{*}_{i=1}^{N}\nu_i(x_i,A)\right]. \tag{2.9}$$

**证** (1) 任取 $A\in\mathscr{E}$, 当 $f=\mathbf{1}_A$ 为 $A$ 上的示性函数时, (2.8) 式显然成立, 再用单调系定理, 易见对任意 $\mathscr{E}$ 可测的有界实值函数 $f$, (2.7) 式恒成立.

(2) 令 $\delta_i(\bullet)=\int_E \mu_i(dx_i)\nu_i(x_i,\bullet)$, 则

$$\begin{aligned}(2.9)\text{ 右边}&=\int_E \mu_1(dx_1)\cdots\int_E \mu_N(dx_N)\\&\quad\left[\int_E \nu_1(x_1,dy_1)\cdots\int_E \nu_n(x_N,dy_N)\right]\bullet\mathbf{1}_A(y_1+\cdots+y_N)\\&=\int_E \delta_1(dy_1)\cdots\int_E \delta(dy_N)\mathbf{1}_A(y_1+\cdots+y_N)\\&=(2.9)\text{ 左边}.\end{aligned}$$

命题 2.2 证毕.

**命题 2.3** 设随机变量 $X$ 的分布函数为 $F(x)$, 则 $F(x)$ 所产生的 (L-S) 测度 $\mu_F$ 就是 $X$ 的分布 $P\circ X^{-1}$.

**证** 对任何实数 $x$, 都有

$$\begin{aligned}\mu_F((-\infty,x])&=F(x)=P(X\leqslant x)\\&=(P\circ X^{-1})((-\infty,x]).\end{aligned}$$

利用单调系定理立得: 对任何 $A\in\mathscr{B}(\mathbf{R})$ 有

$$\mu_F(A)=(P\circ X^{-1})(A).$$

命题 2.3 证毕.

**命题 2.4** $n$ 个相互独立的随机变量 $X_1,\cdots,X_N$ 之和 $X=\sum\limits_{i=1}^{N}X_i$ 的分布 $P\circ X^{-1}$ 等于诸 $X_i$ 的分布 $P\circ X_i^{-1}(i=1,\cdots,N)$ 的卷积.

**证** 设 $F$ 为 $(X_1,\cdots,X_N)$ 的联合分布函数, $F_i$ 为 $X_i$ 的分布函数, $\mu_F$ 和 $\mu_{F_i}$ 分别为 $F$ 和 $F_i$ 所产生的 $(\text{L-S})_N$ 测度和 (L-S) 测度. 由于 $X_1,\cdots,X_N$ 相互独立, 在 (2.3) 中取 $k=n, t_i=i, (i=1,\cdots,N)$ 有 $F(x_1,\cdots,x_N)=\prod\limits_{i=1}^{N}F_i(x_i)$, 从而

$$\mu_F=\mathop{\times}_{i=1}^{N}\mu_{F_i}. \tag{2.10}$$

由卷积的定义及 (2.10) 得: 对任何 $A \in \mathscr{B}(\mathbf{R})$,

$$(P \circ X^{-1})(A) = \int_\Omega P(d\omega)\mathbf{1}_A(X(\omega))$$

$$= \int_\Omega P(d\omega)\mathbf{1}_A(X_1(\omega) + \cdots + X_N(\omega))$$

$$= \int_{\mathbf{R}^N} \mu_F(dx_1, \cdots, dx_N)\mathbf{1}_A(x_1 + \cdots + x_N)$$

$$= \int_{\mathbf{R}} \mu_{F_1}(dx_1) \cdots \int_{\mathbf{R}} \mu_{F_N}(dx_N)\mathbf{1}_A(x_1 + \cdots + x_N)$$

$$= \left(\mathop{*}_{i=1}^{N} \mu_{F_i}\right)(A) = \left(\mathop{*}_{i=1}^{N} P \circ X_i^{-1}\right)(A).$$

命题 2.4 证毕.

## §3 随机变量的矩

随机变量的分布, 是刻画随机变量的一切概率性质的重要工具, 但在具体问题中, 分布往往不易得知. 于是退而求其次, 用随机变量的各种矩, 例如数学期望、方差、协方差、…… 来估计随机变量的一些概率性质.

**定义 3.1** 设 $k$ 为一个正整数, $X$ 和 $Y$ 为概率空间 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量. 如果 $\int_\Omega |X(\omega)|^k P(d\omega) < \infty$, 则称 $X$ 的 $k$ 阶原点矩存在, 称 $\int_\Omega X(\omega)^k P(d\omega)$ 为 $X$ 的 $k$ **阶原点矩**. $X$ 的一阶原点矩 $\int_\Omega X(\omega)P(d\omega)$ 称为 $X$ 的**数学期望** (或者期望, 或者均值), 记之为 $\boldsymbol{E}_P(X)$. 在无混淆的情况下 (即在问题的论述中只有一个概率测度 $P$ 时), 简记 $\boldsymbol{E}_P$ 为 $\boldsymbol{E}$.

当 $\int_\Omega |X(\omega) - \boldsymbol{E}(X)|^k P(d\omega) < \infty$, 称 $X$ 的 $k$ 阶中心矩存在, 而称 $\int_\Omega (X(\omega) - \boldsymbol{E}(X))^k P(d\omega)$ 为 $X$ 的 $k$ **阶中心矩** ($k \geqslant 1$). $X$ 的二阶中心矩称为 $X$ 的**方差**, 记之为 $\mathrm{var}_P(X)$. 当 $\boldsymbol{E}_P(X^2) < \infty, \boldsymbol{E}_P(Y^2) < \infty$, 称 $\boldsymbol{E}_P((X - \boldsymbol{E}_P(X))(Y - \boldsymbol{E}_P(Y)))$ 为 $X, Y$ 的协方差, 记之为 $\mathrm{cov}_P(X, Y)$. 简记 $\mathrm{var}_P(X) = \mathrm{var}(X), \mathrm{cov}_P(X, Y) = \mathrm{cov}(X, Y)$.

**命题 3.1** 设 $X_1, \cdots, X_N$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量, $g$ 是 $\mathbf{R}^N$ 上的实值的 Borel 可测函数, $F$ 是 $(X_1, \cdots, X_N)$ 的联合分布函数, $\mu_F$ 是由 $F$ 产生的 $(\text{L-S})_N$ 测度, 则

$$\int_{\mathbf{R}^N} g(x_1, \cdots, x_N)\mu_F(dx_1, \cdots, dx_N)$$

$$= \int_\Omega g(X_1(\omega), \cdots, X_N(\omega))P(d\omega). \tag{3.1}$$

(上式只要一边有意义, 则另一边也有意义且相等. 当然, 当 $g$ 有界时, (3.1) 左右两边都是有限数且相等; 当 $g \geqslant 0$ 时, (3.1) 左右两边都有意义且相等, 但不一定是有限数.)

**证** (1) 对任意 $A \in \mathscr{B}(\mathbf{R}^N)$, 当 $g(x) = \mathbf{1}_A(x)(x \in \mathbf{R}^N)$ 时, (3.1) 左边 $= \mu_F(A) = P(X^{-1}(A)) =$ (3.1) 右边. (此处 $X = (X_1, \cdots, X_N)$.)

(2) 若 $g \geqslant 0$, 则由 (1), 并用单调系定理可证 (3.1) 式左右两边均有意义 (可能为 $\infty$) 且相等.

(3) 若 $g$ 有界, 则由 (1) 并用单调系定理可证 (3.1) 式左右两边均为实数且相等.

(4) 对一般情形, 令

$$g^+(x) = \begin{cases} g(x), & \text{当 } g(x) > 0, x \in \mathbf{R}^N, \\ 0, & \text{当 } g(x) \leqslant 0, x \in \mathbf{R}^N, \end{cases}$$

$$g^-(x) = \begin{cases} -g(x), & \text{当 } g(x) \leqslant 0, x \in \mathbf{R}^N, \\ 0, & \text{当 } g(x) > 0, x \in \mathbf{R}^N, \end{cases}$$

则 $g = g^+ - g^-, g^+ \geqslant 0, g^- \geqslant 0, g^+$ 和 $g^-$ 皆 Borel 可测. 所以由 (2) 有

$$\begin{aligned} &\int_{\mathbf{R}^N} g^+(x_1, \cdots, x_N)\mu_F(dx_1, \cdots, dx_N) \\ = &\int_\Omega g^+(X_1(\omega), \cdots, X_N(\omega))P(d\omega). \end{aligned} \tag{3.2}$$

类似地, 以 $g^-$ 代 $g^+$, (3.2) 式亦然成立. 命题 3.1 证毕.

命题 3.1 告诉我们: $(\Omega, \mathscr{F}, P)$ 上的抽象积分可以化为相空间 $(\mathbf{R}^N, \mathscr{B}(\mathbf{R}^N), P \circ X^{-1} = \mu_F)$ 上的 L-S 积分, 其中 $F$ 是 $X = (X_1, \cdots, X_N)$ 的联合分布函数, $\mu_F$ 是 $F$ 产生的 $(\text{L-S})_N$ 测度, $P \circ X^{-1}$ 是 $X$ 的分布, 即对任何 $A \in \mathscr{B}(\mathbf{R}^N), (P \circ X^{-1})(A) = P(X^{-1}(A))$.

**命题 3.2** 设 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量 $X$ 和 $Y$ 分别具有分布函数 $F$ 和 $G, H$ 是 $(X, Y)$ 的联合分布函数, $\mu_F, \mu_G$ 和 $\mu_H$ 分别为 $F, G$ 和 $H$ 产生之 (L-S) 测度和 $(\text{L-S})_2$ 测度, 则

$$\boldsymbol{E}(X) = \int_{\mathbf{R}} x\mu_F(dx), \quad \text{当 } \boldsymbol{E}(|X|) < \infty;$$

$$\text{var}(X) = \int_{\mathbf{R}} (x - \boldsymbol{E}(X))^2\mu_F(dx), \quad \text{当 } \boldsymbol{E}(|X|^2) < \infty;$$

$$\text{cov}(X, Y) = \int_{\mathbf{R}^2} (x - \boldsymbol{E}(X))(y - \boldsymbol{E}(Y))\mu_H(dx, dy),$$

当 $\boldsymbol{E}(|X|^2) < \infty$ 且 $\boldsymbol{E}(|Y|^2) < \infty$ 时. 特别地,

(1) 若 $X$ 是离散的且有概率分布 $p_k = P(X = x_k), k = 0, 1, \cdots$; $(X, Y)$ 是离散的且有联合概率分布 $p_{j,k} = P(X = x_j, Y = y_k), j, k = 0, 1, \cdots$,

当 $\sum\limits_{k=0}^{\infty} |x_k| p_k < \infty$ 时, $\boldsymbol{E}(X) = \sum\limits_{k=0}^{\infty} x_k p_k$;

当 $\sum\limits_{k=0}^{\infty} x_k^2 p_k < \infty$ 时, $\mathrm{var}(X) = \sum\limits_{k=0}^{\infty} (x_k - E(X))^2 p_k$;

当 $\sum\limits_{k=0}^{\infty} x_k^2 p_k < \infty, \sum\limits_{k=0}^{\infty} y_k^2 q_k < \infty \left(q_k = \sum\limits_{j=0}^{\infty} p_{j,k}\right)$ 时,

$$\mathrm{cov}(X, Y) = \sum_{j,k=0}^{\infty} p_{j,k}(x_j - E(X))(y_k - E(Y)).$$

(2) 若 $X$ 是连续型的且有密度函数 $f$, $(X, Y)$ 也是连续型的且有联合密度函数 $h$,

当 $\int_{\mathbf{R}} |x| f(x) dx < \infty$ 时, $\boldsymbol{E}(X) = \int_{\mathbf{R}} x f(x) dx$;

当 $\int_{\mathbf{R}} |x|^2 f(x) dx < \infty$ 时, $\mathrm{var}(X) = \int_{\mathbf{R}} (x - \boldsymbol{E}(X))^2 f(x) dx$;

当 $\int_{\mathbf{R}} |x|^2 f(x) dx < \infty$ 且 $\int_{\mathbf{R}} |x|^2 g(x) dx < \infty$ 时 (其中 $g(x) = \int_{\mathbf{R}} h(x, y) dy$),

$$\mathrm{cov}(X, Y) = \int_{\mathbf{R}^2} h(x, y)(x - \boldsymbol{E}(X))(y - \boldsymbol{E}(Y)) dx dy.$$

**证** 利用命题 3.1 及对离散测度 (只在可数个点上有正测度值之测度) 的积分即级数, 就可以证明命题 3.2.

下面我们再列举数学期望与方差的几条简单性质, 设 $X_i$ 皆为实值随机变量, $1 \leqslant i \leqslant N$.

**命题 3.3** (1) 若 $\boldsymbol{E}(X_i)$ 存在, $c_i$ 为实数, $1 \leqslant i \leqslant N$, 则 $\boldsymbol{E}\left(\sum\limits_{i=1}^{N} c_i X_i\right)$ 也存在且等于 $\sum\limits_{i=1}^{N} c_i \boldsymbol{E}(X_i)$;

(2) 若 $\boldsymbol{E}(X_i)$ 存在, $X_1, \cdots, X_N$ 相互独立, $\boldsymbol{E}\left(\prod\limits_{i=1}^{N} X_i\right)$ 也存在, 则

$$\boldsymbol{E}\left(\prod_{i=1}^{N} X_i\right) = \prod_{i=1}^{N} \boldsymbol{E}(X_i);$$

(3) 若 $\mathrm{var}(X)$ 存在, $a$ 为实数, 则

$$\mathrm{var}(a) = 0, \quad \mathrm{var}(aX) = a^2 \mathrm{var}(X);$$

(4) 若 $\mathrm{var}(X_i)$ 存在 $(1 \leqslant i \leqslant N)$, $X_1, \cdots, X_N$ 相互独立, 则

$$\mathrm{var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} \mathrm{var}(X_i).$$

**证** 用 $\boldsymbol{E}(X), \mathrm{var}(X)$ 的定义、积分的线性性质和命题 2.1 立即可证命题 3.3.

**定义 3.2** 设 $X$ 和 $Y$ 皆为二阶原点矩都存在的实值随机变量 (从而其方差 $\mathrm{var}(X), \mathrm{var}(Y)$ 和协方差 $\mathrm{cov}(X,Y)$ 都存在), 若 $\mathrm{var}(X)>0, \mathrm{var}(Y)>0$, 则称

$$\rho(X,Y) \overset{\text{def.}}{=\!=} \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}} \tag{3.3}$$

为 $X$ 和 $Y$ 的相关系数.

**命题 3.4** 设实值随机变量 $X_1,\cdots,X_N$ 的二阶原点矩都存在, 则

$$\mathrm{var}\left(\sum_{i=1}^{N} X_i\right) = \sum_{i=1}^{N}\mathrm{var}(X_i) + \sum_{i\neq j}\mathrm{cov}(X_i,X_j). \tag{3.4}$$

特别地, 若 $X_1$ 与 $X_2$ 相互独立, 则 $\mathrm{cov}(X_1,X_2)=0$, 更特别地, 若还有 $\mathrm{var}(X_i)>0(i=1,2)$, 则 $\rho(X_1,X_2)=0$.

**证** 由 $\mathrm{var}(X_i)$、$\mathrm{cov}(X_1,X_2)$ 和 $\rho(X_1,X_2)$ 的定义直接可证命题 3.4.

**例 3.1** 若随机变量 $X$ 是离散的, 且其概率分布为例 1.1 中的二项分布 $\left\{p_k=\binom{n}{k}p^k(1-p)^{n-k}, k=0,1,\cdots,n\right\}$, 则 $\boldsymbol{E}(X)=np, \mathrm{var}(X)=np(1-p)$.

**例 3.2** 若离散随机变量 $X$ 的概率分布是例 1.2 中的 Poisson 分布 $\left\{p_k= e^{-\lambda}\frac{\lambda^k}{k!}, k=0,1,2,\cdots\right\}$, 则 $\boldsymbol{E}(X)=\mathrm{var}(X)=\lambda$.

**例 3.3** 若连续型的随机变量 $X$ 的密度函数是本章 (1.7) 式中的正态密度函数, 则 $\boldsymbol{E}(X)=a, \mathrm{var}(X)=\sigma^2$.

## §4 随机元及其数学期望

在 §1 和 §3 中, 我们讨论过实值随机变量及其各种矩 (包括数学期望). 在这一节中, 我们将要研究概率空间上取值于一般的可测空间上的随机元.

**定义 4.1** 设 $(\Omega,\mathscr{F},P)$ 是一个概率空间, $(\boldsymbol{B},\mathscr{B})$ 是一个任意的可测空间, 若

$$X:\Omega\mapsto \boldsymbol{B}$$

满足 $X^{-1}(A)\in\mathscr{F}$ 对任意 $A\in\mathscr{B}$ 成立, 即是 $X$ 是 $\mathscr{F}/\mathscr{B}$ 可测映射, 记之为 $X\in\mathscr{F}/\mathscr{B}$, 则称 $X$ 是 $\Omega$ 上的一个 $\boldsymbol{B}$ **值随机元**, 简称 $\boldsymbol{B}$ 值随机元或**随机元**.

前三节所讨论的实值随机变量, 实为 $\mathbf{R}$ 值随机元, 或 $\mathscr{F}/\mathscr{B}(\mathbf{R})$ 可测映射.

从实值随机变量到 $\boldsymbol{B}$ 值随机元, 这是本质性的推广. 实值随机变量的一些性质, 对一般的 $\boldsymbol{B}$ 值随机元是否仍然保持是需要验证的.

**命题 4.1** 设 $(\boldsymbol{B},\mathscr{B},d)$ 是一个可测距离空间, 即 $d$ 是 $\boldsymbol{B}$ 上的一个距离, $\mathscr{B}$ 是由 $\boldsymbol{B}$ 中的全体开集所产生的 Borel $\sigma$ 代数. $\{X_n, n\geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的一列 $\boldsymbol{B}$ 值随机元. 如果对任何 $\omega\in\Omega$, 都有

$$\lim_{n\to\infty} X_n(\omega)=X(\omega)\ 存在,$$

则 $X$ 也是一个 $\boldsymbol{B}$ 值随机元.

**证** 记 $\boldsymbol{B}$ 中全体闭集为 $\mathfrak{M}^{\circ}$. 任取 $F\in\mathfrak{M}$, 令

$$f(x)=\inf\{d(x,y):y\in F\},\quad (x\in\boldsymbol{B}),\tag{4.1}$$

则 $f(x)$ 是 $\boldsymbol{B}$ 上的实值连续函数 (参见文献 [35] 第一章定理 6.1) 而且 $F=\{x\in\boldsymbol{B}:f(x)=0\}$. 所以 $X^{-1}(F)=\{\omega\in\Omega:f(X(\omega))=0\}$. 而由复合函数的可测性及 $f$ 是连续函数得知 $f(X_n(\omega))$ 是可测空间 $(\Omega,\mathscr{F})$ 到可测空间 $(\mathbf{R},\mathscr{B}(\mathbf{R}))$ 的可测函数, 所以 $f(X(\omega))=\lim\limits_{n\to\infty}f(X_n(\omega))$ 亦然. 因此, $X^{-1}(F)=\{\omega\in\Omega:f(X(\omega))=0\}\in\mathscr{F}$, 从而

$$X^{-1}(\mathscr{B})=X^{-1}(\sigma(\mathfrak{M}))=\sigma(X^{-1}(\mathfrak{M}))\subset\mathscr{F}.$$

命题 4.1 证毕.

**命题 4.2** 设 $(\boldsymbol{B},\mathscr{B},d)$ 是可分的可测距离空间, 则 $X:\Omega\to\boldsymbol{B}$ 为 $\boldsymbol{B}$ 值随机元的充分必要条件是: 存在一列从 $\Omega$ 到 $\boldsymbol{B}$ 的可数值的 $\boldsymbol{B}$ 值随机元 $\{X_n,n\geqslant 0\}$, 使

$$\lim_{n\to\infty}X_n(\omega)=X(\omega),\quad (对\ \omega\in\Omega\ 一致成立).\tag{4.2}$$

**证** 充分性. 由命题 4.1 立即可得.

必要性. 由 $(\boldsymbol{B},d)$ 的可分性, 可取一个在 $\boldsymbol{B}$ 中稠密的可数子集 $\{x_n,n\geqslant 1\}$. 对每个 $n\geqslant 1$, 定义

$$A_{n,i}=\{x\in\boldsymbol{B}:d(x,x_i)\leqslant\frac{1}{n}\}-\bigcup_{j=1}^{i-1}A_{n,j},\quad i=1,2,\cdots,\ 约定\ \bigcup_{j=1}^{0}A_{n,j}=\varnothing.$$

于是 $\{A_{n,i},i=1,2,\cdots\}$ 两两不交且其并为全空间 $\boldsymbol{B}$. 再对每个 $n\geqslant 1$ 定义

$$X_n(\omega)=\sum_{i=1}^{\infty}x_i\mathbf{1}_{X^{-1}(A_{n,i})}(\omega)\quad(\omega\in\Omega),\tag{4.3}$$

易见 $X_n$ 是从 $\Omega$ 到 $\boldsymbol{B}$ 的可数值的 $\boldsymbol{B}$ 值随机元, 且有

$$d(X_n(\omega), X(\omega)) \leqslant \frac{1}{n} \quad (\forall \omega \in \Omega, n \geqslant 1).$$

故 $\{X_n\}$ 即为所求. 命题 4.2 证毕.

**命题 4.3** 设 $(\boldsymbol{B}, \mathscr{B}, d)$ 是可分的可测距离空间, $X$ 和 $Y$ 是 $(\Omega, \mathscr{F}, P)$ 上的 $\boldsymbol{B}$ 值随机元, 则 $d(X(\omega), Y(\omega))$ 是 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量.

**证** 定义

$$Z(\omega) = (X(\omega), Y(\omega)),$$

则对任何 $A_1, A_2 \in \mathscr{B}(\boldsymbol{B})$, 有

$$Z^{-1}(A_1 \times A_2) = X^{-1}(A_1) \bigcap Y^{-1}(A_2) \in \mathscr{F}.$$

利用单调系定理由上式可证: 对任何 $c \in \mathscr{B}(\boldsymbol{B}) \times \mathscr{B}(\boldsymbol{B})$, 有 $Z^{-1}(c) \in \mathscr{F}$. 如果能证明 $\mathscr{B}(\boldsymbol{B}) \times \mathscr{B}(\boldsymbol{B}) = \mathscr{B}(\boldsymbol{B} \times \boldsymbol{B})$, 则 $Z$ 关于 $\mathscr{F}/\mathscr{B}(\boldsymbol{B} \times \boldsymbol{B})$ 可测. 但是 $d(x, y)$ 是从 $\boldsymbol{B} \times \boldsymbol{B}$ 到 $\mathbf{R}$ 的连续映射. 从而是 $\mathscr{B}(\boldsymbol{B} \times \boldsymbol{B})/\mathscr{B}(\mathbf{R})$ 可测映射. 利用复合函数的可测性定理, $d(X, Y)$ 是由 $\Omega$ 到 $\mathbf{R}$ 的实值随机变量. 下面补证 $\mathscr{B}(\boldsymbol{B}) \times \mathscr{B}(\boldsymbol{B}) = \mathscr{B}(\boldsymbol{B} \times \boldsymbol{B})$.

**引理 4.1** 设 $(\boldsymbol{B}_1, d_1), (\boldsymbol{B}_2, d_2)$ 都是可分的距离空间, 则乘积距离空间 $(\boldsymbol{B}_1 \times \boldsymbol{B}_2, d_1 \times d_2)$ 上的 Borel $\sigma$ 代数 $\mathscr{B}(\boldsymbol{B}_1 \times \boldsymbol{B}_2) = \mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2)$.

**证** 令 $f_i$ 是由 $\boldsymbol{B}_1 \times \boldsymbol{B}_2$ 到 $\boldsymbol{B}_i$ 的投影算子, $i = 1, 2$, 则 $f_i$ 是由 $(\boldsymbol{B}_1 \times \boldsymbol{B}_2, d_1 \times d_2)$ 到 $(\boldsymbol{B}_i, d_i)$ 的连续映射, 更有 $f_i^{-1}(\mathscr{B}(\boldsymbol{B}_i)) \subset \mathscr{B}(\boldsymbol{B}_1 \times \boldsymbol{B}_2), i = 1, 2$. 所以对任何 $A_i \in \mathscr{B}(\boldsymbol{B}_i)$ 有

$$A_1 \times A_2 = f_1^{-1}(A_1) \bigcap f_2^{-1}(A_2) \in \mathscr{B}(\boldsymbol{B}_1 \times \boldsymbol{B}_2), \tag{4.4}$$

从而

$$\begin{aligned} \mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2) &= \sigma(\{A_1 \times A_2 : A_i \in \mathscr{B}(\boldsymbol{B}_i), i = 1, 2\}) \\ &\subset \mathscr{B}(\boldsymbol{B}_1 \times \boldsymbol{B}_2). \end{aligned}$$

$$\text{再证 } \mathscr{B}(\boldsymbol{B}_1 \times \boldsymbol{B}_2) \subset \mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2). \tag{4.5}$$

为此, 只需证 $\boldsymbol{B}_1 \times \boldsymbol{B}_2$ 中任一开集均属于 $\mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2)$. 由于 $(\boldsymbol{B}_i, d_i)$ 是可分的, 所以有可数的拓扑基 $\{U_n^i, n = 1, 2, \cdots\}$. 而柱集系 $\{f_i^{-1}(U_n^i), n = 1, 2, \cdots\}(i = 1, 2)$ 是 $\boldsymbol{B}_1 \times \boldsymbol{B}_2$ 上的拓扑的一次基底. 而这种柱集和它们的有限交都属于 $\mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2)$. 因此作为它们的可数并的任一开集也属于 $\mathscr{B}(\boldsymbol{B}_1) \times \mathscr{B}(\boldsymbol{B}_2)$. 引理 4.1 证毕.

**引理 4.2** 补证了命题 4.3 的特征的一半. 命题 4.3 证毕.

**附注 4.1** 命题 4.3 中 $(\boldsymbol{B},\mathscr{B},d)$ 的可分性条件是不可去的, 因为有反例说明, 当 $(\boldsymbol{B},\mathscr{B},d)$ 不可分时, $d(X(\omega),Y(\omega))$ 不是 $(\Omega,\mathscr{F},P)$ 的实值随机变量.(参见文献 [111] P.4.)

**定理 4.1**[111] 设 $(\boldsymbol{B},\|\bullet\|)$ 是可分的 Banach 空间, $\mathscr{B}(\boldsymbol{B})$ 是 $\boldsymbol{B}$ 上的 Borel $\sigma$ 代数, 则 $X$ 是 $(\Omega,\mathscr{F},P)$ 上的 $\boldsymbol{B}$ 值随机元的充分必要条件是: 对 $\boldsymbol{B}$ 上的任一有界线性泛函 $f, f(X)$ 都是 $(\Omega,\mathscr{F},P)$ 上的实值随机变量.

**证** 必要性显然成立 (因为连续泛函必可测.)

充分性. 令 $\boldsymbol{B}^*$ 是定义在 $\boldsymbol{B}$ 上的有界线性泛函全体, 再令

$$\mathscr{E}=\{f^{-1}(A): f\in\boldsymbol{B}^*, A\in\mathscr{B}(\mathbf{R})\}. \tag{4.6}$$

由连续泛函必可测知: $\mathscr{E}\subset\mathscr{B}(\boldsymbol{B})$. 再任取 $f\in\boldsymbol{B}^*, A\in\mathscr{B}(\mathbf{R})$, 由 $f(X)$ 是实值随机变量知:

$$X^{-1}(f^{-1}(A))=(f(X))^{-1}(A)\in\mathscr{F}. \tag{4.7}$$

即

$$X^{-1}(\mathscr{E})\subset\mathscr{F}. \tag{4.8}$$

若能证

$$\sigma(\mathscr{E})=\mathscr{B}(\boldsymbol{B}), \tag{4.9}$$

则由第一章引理 1.1 得

$$X^{-1}(\mathscr{B}(\boldsymbol{B}))=X^{-1}(\sigma(\mathscr{E}))=\sigma(X^{-1}(\mathscr{E}))\subset\mathscr{F}, \tag{4.10}$$

即 $X$ 是 $\boldsymbol{B}$ 值随机元.

现在补证 (4.9). 而前面已证 $\mathscr{E}\subset\mathscr{B}(\boldsymbol{B})$, 从而 $\sigma(\mathscr{E})\subset\mathscr{B}(\boldsymbol{B})$, 所以要证 (4.9), 只需证明 $\mathscr{B}(\boldsymbol{B})\subset\sigma(\mathscr{E})$. 由于 $\boldsymbol{B}$ 是可分的, 为此, 又只需证明 $\boldsymbol{B}$ 中的任何一个闭球

$$B(x,r)\overset{\text{def.}}{=\!=}\{y\in\boldsymbol{B}:\|x-y\|\leqslant r\}\in\sigma(\mathscr{E}),(r\in\mathbf{R},x\in\boldsymbol{B}).$$

设 $\{x_n\}$ 是 $\boldsymbol{B}$ 中一个可数稠子集. 对每个 $n\geqslant 1$, 取 $f_n\in\boldsymbol{B}^*$, 使 $\|f_n\|=1, f_n(x_n-x)=\|x_n-x\|$. (由 Hahn-Banach 定理, 这样的 $f_n$ 是存在的.) 再令

$$B_1=\{y\in\boldsymbol{B}: f_n(y-x)\leqslant r, n=1,2,\cdots\}, \tag{4.11}$$

则 $B_1\in\sigma(\mathscr{E})$, 且 $B_1\supset B(x,r)$.

若 $y\overline{\in}B(x,r)$, 则 $\|x-y\|>r$. 取 $x_k$ 使

$$\|y-x_k\|<\frac{1}{2}(\|y-x\|-r),$$

则

$$\begin{aligned}\|x_k-x\| &\geqslant \|y-x\|-\|y-x_k\| \\ &> \|y-x\|-\frac{1}{2}(\|y-x\|-r) \\ &= \frac{1}{2}(\|y-x\|-r); \\ |f_k(y-x)-\|x_k-x\|| &= |f_k(y-x)-f_k(x_k-x)| \\ &\leqslant \|y-x_k\|<\frac{1}{2}(\|y-x\|-r),\end{aligned}$$

从而

$$\begin{aligned}f_k(y-x) &= \|x_k-x\|-(\|x_k-x\|-f_k(y-x)) \\ &> \frac{1}{2}(\|y-x\|+r)-\frac{1}{2}(\|y-x\|-r) \\ &= r,\end{aligned}$$

此即 $y\overline{\in}B_1$. 总之 $B(x,r)=B_1\in\sigma(\mathscr{E})$. 定理 4.1 证毕.

**附注 4.2** 定理 4.1 可以改写成下列更一般的形式:

设 $(\boldsymbol{B},\|\bullet\|)$ 是可分的 Banach 空间, $(\Omega,\mathscr{F})$ 是任一可测空间, $X:\Omega\mapsto\boldsymbol{B}$, 则 $X$ 是 $\mathscr{F}/\mathscr{B}(\boldsymbol{B})$ 可测的充分必要条件是: 对 $\boldsymbol{B}$ 上的任何一个有界线性泛函 $f$ (即 $f\in\boldsymbol{B}^*$), $f(X)$ 都是 $\mathscr{F}/\mathscr{B}(\mathbf{R})$ 可测的.

因为定理 4.1 的证明中, 并未要求 $\sigma$ 代数 $\mathscr{F}$ 上有任何测度.

以后如不声明, Banach 空间中的极限都是强极限, 即是依范数收敛的极限.

**附注 4.3** 设 $(E_i,\mathscr{E}_i)(i=1,2)$ 是两个可测空间, $f:E_1\mapsto E_2$, 若 $f^{-1}(\mathscr{E}_2)\subset\mathscr{E}_1$, 则称 $f$ 关于 $\mathscr{E}_1$ 和 $\mathscr{E}_2$ 可测, 或称之为 $f$ 关于 $\mathscr{E}_1/\mathscr{E}_2$ 可测, 或记之为 $f\in\mathscr{E}_1/\mathscr{E}_2$.

**推论 4.1** 设 $(\boldsymbol{B},\|\bullet\|)$ 是可分的 Banach 空间, $X_i$ 是 $(\Omega,\mathscr{F},P)$ 上的 $\boldsymbol{B}$ 值随机元, $\alpha_i$ 为实数, $i=1,2,\cdots,n$, 则 $\sum\limits_{i=1}^{n}\alpha_iX_i$ 也是 $(\Omega,\mathscr{F},P)$ 上的 $\boldsymbol{B}$ 值随机元.

**定义 4.2** 设 $(\Omega,\mathscr{F})$ 是任一可测空间 $\{A_1,\cdots,A_n\}\subset\mathscr{F}$, $\{A_i\}$ 两两不交且 $\bigcup\limits_{i=1}^{n}A_i=\Omega$, 则称 $\{A_i\}$ 是 $\Omega$ 的一个有限的可测分划. 设 $(\boldsymbol{B},\|\bullet\|)$ 是一个 Banach 空间, $\{c_1,\cdots,c_n\}\subset\boldsymbol{B}$, $\{A_1,\cdots,A_n\}$ 是 $\Omega$ 的一个有限的可测分划, 则称

$$X=\sum_{i=1}^{n}c_i\mathbf{1}_{A_i} \tag{4.12}$$

为 $(\Omega,\mathscr{F})$ 上的一个简单的 $\boldsymbol{B}$ 值函数. 特别地若 $\mathscr{F}$ 上还有一个完备测度 $\mu$, 则称 $(\Omega,\mathscr{F},\mu)$ 上的 $\boldsymbol{B}$ 值函数 $X$ 是 $\mathscr{F}$ 强可测的 $\boldsymbol{B}$ 值函数, 简称强可测的 $\boldsymbol{B}$ 值函数, 如果存在一列简单的 $\boldsymbol{B}$ 值函数

$$\left\{X_n=\sum_{i=1}^{k_n}c_i^{(n)},\mathbf{1}_{A_i^{(n)}},n\geqslant 1\right\},$$

使

$$X=\lim_{n\to\infty}X_n=\lim_{n\to\infty}\sum_{i=1}^{k_n}c_i^{(n)}\mathbf{1}_{A_i^{(n)}}\mu-\text{a.s.} \tag{4.13}$$

其中 $\{c_1^{(n)},\cdots,C_{k_n}^{(n)}\}\subset\boldsymbol{B},\{A_1^{(n)},\cdots,A_{k_n}^{(n)}\}$ 是 $\Omega$ 的有限的可测分划, $(n\geqslant 1)$. (4.13) 中的极限是强极限, 即依范数 $\|\bullet\|$ 收敛的极限.

特别地, 若 $\mu=P$ 是完备的概率测度, 则 $(\Omega,\mathscr{F},P)$ 上的简单 $\boldsymbol{B}$ 值函数称为简单的 $\boldsymbol{B}$ 值随机元. 显然, 简单的 $\boldsymbol{B}$ 值随机元是 $\boldsymbol{B}$ 值随机元. $(\Omega,\mathscr{F},P)$ 上的强可测 $\boldsymbol{B}$ 值函数称强 $\boldsymbol{B}$ 值随机元, 显然强 $\boldsymbol{B}$ 值随机元是 $\boldsymbol{B}$ 值随机元.

以后如不声明, Banach 空间中的极限都是强极限, 即是依范数收敛的极限.

**命题 4.4** 设 $(\Omega,\mathscr{F},\mu)$ 是完备测度空间, $(\boldsymbol{B},\|\bullet\|)$ 是 Banach 空间, $X$ 是 $\mathscr{F}$ 强可测 $\boldsymbol{B}$ 值函数, 则

(1) $\|X\|:\Omega\mapsto\mathbf{R}$ 是 $\mathscr{F}/\mathscr{B}(\mathbf{R})$ 可测的实值函数;

(2) 存在一列简单的 $\boldsymbol{B}$ 值函数 $\{X_n\}$ 使

$$\|X_n(\omega)\|\leqslant 2\|X(\omega)\|\quad(\forall\omega\in\Omega),$$

且

$$\lim_{n\to\infty}X_n=X\quad\mu-\text{a.s.}.$$

证明甚易, 读者可作为习题验证之.

**定理 4.2** 设 $(\Omega,\mathscr{F},\mu)$ 是完备有限的测度空间, $(\boldsymbol{B},\|\bullet\|)$ 是可分 Banach 空间, 则下列陈述等价:

(1) $X$ 是 $\mathscr{F}/\mathscr{B}(\boldsymbol{B})$ 可测的 $\boldsymbol{B}$ 值函数;

(2) 存在 $\mathscr{F}/\mathscr{B}(\boldsymbol{B})$ 可测的可数值的 $\boldsymbol{B}$ 值函数列 $\{X_n=\sum\limits_{i=1}^{\infty}x_i\mathbf{1}_{A_i^{(n)}},n\geqslant 1\}$, 使

$$\lim_{n\to\infty}X_n(\omega)=X(\omega)(\text{在 }\omega\in\Omega\text{ 上一致成立}); \tag{4.14}$$

(3) 存在 $\mathscr{F}/\mathscr{B}(\boldsymbol{B})$ 可测的可数值的 $\boldsymbol{B}$ 值函数列 $\{X_n=\sum\limits_{i=1}^{\infty}x_i\mathbf{1}_{A_i^{(n)}},n\geqslant 1\}$, 使

$$\lim_{n\to\infty}X_n(\omega)=X(\omega)(\forall\omega\in\Omega); \tag{4.15}$$

(4) 存在简单的 $\boldsymbol{B}$ 值函数列 $\{X'_n\}$, 使

$$\lim_{n\to\infty} X'_n = X, \quad \mu-\text{a.s.}; \tag{4.16}$$

(5) $X$ 是强可测的 $\boldsymbol{B}$ 值函数;

(6) 对任何 $f \in \boldsymbol{B}^*$ ($\boldsymbol{B}^*$ 之定义见定理 4.1), $f(X)$ 是 $\mathscr{F}/\mathscr{B}(\mathbf{R})$ 可测的.

**证** (1) $\Rightarrow$ (2). 由命题 4.2 立即可得.

(2) $\Rightarrow$ (3). 显然成立.

(3) $\Rightarrow$ (4). 设 (3) 成立. 则存在

$$X_n = \sum_{i=1}^{\infty} x_i \mathbf{1}_{A_i^{(n)}}, \quad (n \geqslant 1).$$

($\{A_i^{(n)}, i \geqslant 1\}$ 是 $\Omega$ 的一个可测分划, $\{x_i, i \geqslant 1\} \subset \boldsymbol{B}(n \geqslant 1)$), 使 (4.15) 成立. 由于 $\mu\left(\bigcup_{i=1}^{\infty} A_i^{(n)}\right) = \sum_{i=1}^{\infty} \mu(A_i^{(n)}) = K < \infty (n \geqslant 1)$, 所以对每个 $n \geqslant 1$, 存在正整数 $k_n$, 使

$$\sum_{i>k_n} \mu(A_i^{(n)}) = \mu\left(\bigcup_{i>k_n} A_i^{(n)}\right) \leqslant \frac{1}{n}. \tag{4.17}$$

取

$$X_n^* = X_n \sum_{i=1}^{k_n} \mathbf{1}_{A_i^{(n)}} = \sum_{i=1}^{k_n} x_i \mathbf{1}_{A_i^{(n)}}, \quad (n \geqslant 1), \tag{4.18}$$

由 $\lim\limits_{n\to\infty} \|X_n(\omega) - X(\omega)\| = 0$ 得 $\lim\limits_{n\to\infty} \|X_n(\omega)\| = \|X(\omega)\|$, 从而

$$\sup_{n\geqslant 1} \|X_n(\omega)\| \leqslant M(\omega) < \infty, \tag{4.19}$$

所以

$$\begin{aligned} \|X_n^*(\omega) - X_n(\omega)\| &= \|X_n(\omega) \sum_{i>k_n} \mathbf{1}_{A_i^{(n)}}(\omega)\| \\ &\leqslant M(\omega) \sum_{i>k_n} \mathbf{1}_{A_i^{(n)}}(\omega) \xrightarrow{(n\to\infty)} 0 \ \left(\text{当 } \omega \in \bigcup_{n=1}^{\infty} \bigcup_{i\leqslant k_n} A_i^{(n)}\right). \end{aligned} \tag{4.20}$$

由 (4.17)~(4.20) 和 $(\Omega, \mathscr{F}, \mu)$ 的完备性, 若令

$$X'_n(\omega) = \begin{cases} X_n^*(\omega), & \text{当 } \omega \in \bigcup\limits_{n=1}^{\infty} \bigcup\limits_{i\leqslant k_n} A_i^{(n)}, \\ x_0, & \text{反之}, \end{cases} \tag{4.21}$$

则 $\{X'_n, n \geqslant 1\}$ 是简单的 $\boldsymbol{B}$ 值函数列, 且使 (4.15) 成立. (注意: 由 $X_n^*$ 的定义 (2.18) 知 $X_n^*$ 是简单的 $\boldsymbol{B}$ 值函数, 从而 $X'_n$ 亦然.)

(4) ⇒ (5). 由强可测的 $\boldsymbol{B}$ 值函数的定义可得.

(5) ⇒ (6). 用附注 4.2 及 $(\Omega, \mathscr{F}, \mu)$ 的完备性立即可证得此结论.

(6) ⇒ (1). 用附注 4.2 立即可得此结论.

**附注 4.4** 定理 4.2 还可以用随机元的语言叙述如下:

**定理 4.3** 设 $(\boldsymbol{B}, \|\bullet\|)$ 是可分的 Banach 空间, $(\Omega, \mathscr{F}, P)$ 是完备的概率空间, $\boldsymbol{B}^*$ 是 $\boldsymbol{B}$ 的对偶空间, 则下列陈述等价:

(1) $X$ 是 $\boldsymbol{B}$ 值随机元;

(2) $X$ 可表为

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad (\text{在 } \omega \in \Omega \text{ 一致, 或逐点成立}),$$

其中

$$X_n = \sum_{i=1}^{\infty} x_i \mathbf{1}_{A_i^{(n)}} \text{ 为可数值 } \boldsymbol{B} \text{ 值随机元};$$

(3) $X$ 是 $\boldsymbol{B}$ 值强随机元;

(4) 对任何 $f \in \boldsymbol{B}^*, f(X)$ 是实值随机变量.

在研究随机元的数学期望以前, 我们要给出 Bochner 积分 (强积分) 的定义及简单性质.

**定义 4.3** 设 $(\Omega, \mathscr{F}, \mu)$ 是完备测度空间, $(\boldsymbol{B}, \|\bullet\|)$ 是 Banach 空间, $X : \Omega \mapsto \boldsymbol{B}$ 是 $\boldsymbol{B}$ 值强可测函数, 且 $\int_\Omega \|X(\omega)\|\mu(d\omega) < \infty$.

(1) 若 $X$ 是简单 $\boldsymbol{B}$ 值函数:

$$X = \sum_{i=1}^{n} x_i \mathbf{1}_{A_i}, \{x_1, \cdots, x_n\} \subset \boldsymbol{B},$$

$\{A_1, \cdots, A_n\}$ 是 $\Omega$ 的一个可测分划, 则定义 $X$ 在 $\Omega$ 上的 Bochner 积分为 $\sum\limits_{i=1}^{n} x_i \mu(A_i)$, 记为

$$(\mathrm{s}) \int_\Omega X(\omega)\mu(d\omega) \stackrel{\text{def.}}{=\!=} (\mathrm{s}) \int_\Omega X d\mu = \sum_{i=1}^{n} x_i \mu(A_i).$$

(2) 对满足定义 4.3 中的条件的 $X$, 由命题 4.4, 总可以取简单 $\boldsymbol{B}$ 值函数列 $\{X_n\}$ 使

$$\|X_n(\omega)\| \leqslant 2\|X(\omega)\| \quad (\forall \omega \in \Omega),$$

$$\lim_{n\to\infty} X_n = X, \quad \mu - \text{a.s.},$$

这时, 定义 $X$ 在 $\Omega$ 上的 **Bochner** 积分为

$$(\mathrm{s})\int_\Omega X d\mu = \lim_{n\to\infty}\left((\mathrm{s})\int_\Omega X_n d\mu\right).$$

为了证明定义 4.3 的合理性, 有必要证明

**定理 4.4** 设 $(\Omega,\mathscr{F},\mu)$ 是完备测度空间, $(\boldsymbol{B},\|\bullet\|)$ 是 Banach 空间, $X:\Omega\mapsto \boldsymbol{B}$ 且 $X$ 是 $\mathscr{F}$ 强可测的. 若 $\int_\Omega\|X(\omega)\|\mu(d\omega)<\infty$, 则对任何简单的 $\boldsymbol{B}$ 值函数列 $\{X_n,n\geqslant 1\}$, 只要满足

$$\lim_{n\to\infty} X_n = X,\quad \mu-\text{a.s.},\tag{4.22}$$

$$\|X_n(\omega)\|\leqslant 2\|X(\omega)\|\quad (\forall\omega\in\Omega),\tag{4.23}$$

必有:

(1) 存在 $x\in\boldsymbol{B}$, 使

$$\lim_{n\to\infty}(\mathrm{s})\int_\Omega X_n d\mu = x;$$

(2)

$$\lim_{n\to\infty}\int_\Omega \|X(\omega)-X_n(\omega)\|\mu(d\omega)=0.\tag{4.24}$$

若还有一列简单的 $\boldsymbol{B}$ 值函数 $\{\widetilde{X}_n,n\geqslant 1\}$ 亦满足 (4.22) 和 (4.23), 则

$$\lim_{n\to\infty}(\mathrm{s})\int_\Omega X_n d\mu = \lim_{n\to\infty}(\mathrm{s})\int_\Omega \widetilde{X}_n d\mu.\tag{4.25}$$

**证** (1) 由命题 4.4 知必存在一列简单的 $\boldsymbol{B}$ 值函数 $\{X_n,n\geqslant 1\}$ 使 (4.22) 和 (4.23) 成立. 由 (4.22), (4.23) 和控制收敛定理得:

$$\begin{aligned}
&\lim_{m,n\to\infty}\left\|(\mathrm{s})\int_\Omega X_m d\mu-(\mathrm{s})\int_\Omega X_n d\mu\right\|\\
\leqslant&\lim_{m,n\to\infty}\int_\Omega\|X_m-X_n\|d\mu\\
=&\int_\Omega\lim_{m,n\to\infty}\|X_m-X_n\|d\mu=0.
\end{aligned}$$

由 Banach 空间的完备性得知: 必存在 $x\in\boldsymbol{B}$ 使

$$\lim_{n\to\infty}\int_\Omega X_n d\mu = x.\tag{4.26}$$

(2) (4.24) 可由 $\{X_n,n\geqslant 1\}$ 之选取及控制收敛定理而证得.

若还有一列简单的 $\boldsymbol{B}$ 值函数 $\{\widetilde{X}_n,n\geqslant 1\}$ 亦满足 (4.22) 和 (4.23), 令

$$Y_n=\begin{cases}X_n, & 当\ n\ 为奇数时,\\ \widetilde{X}_n, & 当\ n\ 为偶数时,\end{cases}$$

则简单的 $\boldsymbol{B}$ 值函数列 $\{Y_n, n \geqslant 1\}$ 亦满足 (4.22) 和 (4.23), 所以必存在 $y \in \boldsymbol{B}$, 使

$$y = \lim_{n\to\infty} (\mathrm{s}) \int_{\Omega} Y_n d\mu,$$

但 $\left\{(\mathrm{s}) \int_{\Omega} X_n d\mu, n \geqslant 1\right\}$ 和 $\left\{(\mathrm{s}) \int_{\Omega} \widetilde{X}_n d\mu, n \geqslant 1\right\}$ 都是 $\left\{(\mathrm{s}) \int_{\Omega} Y_n d\mu, n \geqslant 1\right\}$ 的子序列, 所以

$$\lim_{n\to\infty} (\mathrm{s}) \int_{\Omega} \widetilde{X}_n d\mu = \lim_{n\to\infty} (\mathrm{s}) \int_{\Omega} X_n d\mu = x.$$

定理 4.4 证毕.

**定义 4.4** 设 $(\Omega, \mathscr{F}, \mu)$ 是完备测度空间, $(\boldsymbol{B}, \|\bullet\|)$ 是 Banach 空间, $X$ 是 $\boldsymbol{B}$ 值的 $\mathscr{F}$ 强可测函数, 若定义 4.3 中所定义的

$$(\mathrm{s}) \int_{\Omega} X d\mu$$

在 $\boldsymbol{B}$ 中存在且唯一, 则称 $X$ (在 $\Omega$ 上) 是 Bochner 可积的.

**推论 4.2** $\boldsymbol{B}$ 值 $\mathscr{F}$ 强可测函数 $X$ 是 Bochner 可积的充分必要条件是

$$\int_{\Omega} \|X(\omega)\| \mu(d\omega) < \infty. \tag{4.27}$$

**证** 应用定理 4.4, 可直接得推论 4.2.

**定义 4.5** 与实值可测函数的 Lebesgue 积分类似, 对任何 $A \in \mathscr{F}$, 定义 $\boldsymbol{B}$ 值 $\mathscr{F}$ 强可测函数 $X$ 在 $A$ 上的 Bochner 积分为

$$(\mathrm{s}) \int_{A} X d\mu = (\mathrm{s}) \int_{\Omega} \mathbf{1}_A X d\mu. \tag{4.28}$$

**定理 4.5** 设 $(\Omega, \mathscr{F}, \mu)$ 是完备测度空间, $(\boldsymbol{B}, \|\bullet\|)$ 是 Banach 空间, $X_i$ 是 $\boldsymbol{B}$ 值 $\mathscr{F}$ 强可测函数, $c_i \in \mathbf{R}$, $X_i$ 是 Bochner 可积的, $(i = 1, 2, \cdots, n)$, 则 $\sum\limits_{i=1}^{n} c_i X_i$ 也是 Bochner 可积的, 且

$$(\mathrm{s}) \int_{\Omega} \left(\sum_{i=1}^{n} c_i X_i\right) d\mu = \sum_{i=1}^{n} c_i (\mathrm{s}) \int_{\Omega} X_i d\mu; \tag{4.29}$$

$$\left\|(\mathrm{s}) \int_{\Omega} X_i d\mu\right\| \leqslant \int_{\Omega} \|X_i\| d\mu. \tag{4.30}$$

此外, 对任何 $x \in \boldsymbol{B}$, 任何 $\Omega$ 上的实值可测的 Lebesgue 可积的函数 $u(\omega)$, 恒有

$$(\mathrm{s}) \int_{\Omega} u(\omega) x \mu(d\omega) = \left(\int_{\Omega} u(\omega) \mu(d\omega)\right) x. \tag{4.31}$$

证明甚易, 读者可作为习题验证之.

**定义 4.6** 设 $(\Omega, \mathscr{F}, P)$ 是完备的概率空间, $(\boldsymbol{B}, \|\bullet\|)$ 是 Banach 空间, $X$ 是 $\Omega$ 上的 $\boldsymbol{B}$ 值随机元且 $\mathscr{F}$ 强可测. 若 $\int_\Omega \|X(\omega)\| P(d\omega) < \infty$, 则由推论 4.2, $X$ 是 Bochner 可积的, 定义 $X$ 的 Bochner 积分

$$(\mathrm{s}) \int_\Omega X d\mu$$

为 $X$ 的强 (或 Bochner 意义下) 数学期望, 简称数学期望 (期望, 或均值), 仍用 $\boldsymbol{E}(X)$ 记之. 这时, 称 $X$ 的数学期望 $\boldsymbol{E}(X)$ 存在.

由定理 4.5 可知: 若 $X_i$ 的数学期望 $\boldsymbol{E}(X_i)$ 存在,$c_i$ 是实常数, $i = 1, 2, \cdots, n$, 则 $\sum\limits_{i=1}^n c_i X_i$ 的数学期望 $\boldsymbol{E}\left(\sum\limits_{i=1}^n c_i X_i\right)$ 也存在, 且

$$\boldsymbol{E}\left(\sum_{i=1}^n c_i X_i\right) = \sum_{i=1}^n c_i \boldsymbol{E}(X_i). \tag{4.32}$$

## §5 实值随机变量的条件期望

在初等概率论中, 我们曾经论及过两个随机事件 (可测集合) 之间的条件概率和两个随机变量之间的条件期望, 并且发现过它们在初等概率论中的重要性.

现在我们要研究随机事件 (可测集合) 关于给定的 $\sigma$ 代数的条件概率和随机变量关于给定的 $\sigma$ 代数的条件期望. 这些概念, 在高等概率论中, 特别是在本书后面要研究的 Markov 过程和鞅论中, 是时时不能离开的工具.

**定义 5.1** 设 $X$ 是完备概率空间 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量且 $\boldsymbol{E}(|X|) < \infty$. 再设 $\mathscr{G} \subset \mathscr{F}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数. 由 Radon – Nikodym 定理, 存在一个 $\mathscr{G}/\mathscr{B}(\mathbf{R})$ 可测的实值随机变量 $Y$, 使得

$$\int_A Y(\omega) P(d\omega) = \int_A X(\omega) P(d\omega) \quad (\forall A \in \mathscr{G}). \tag{5.1}$$

不仅如此, 除一个 $P-$ 零测集以外, $Y$ 是唯一决定的. 称 $Y$ 是 $X$ 关于 $\sigma$ 代数 $\mathscr{G}$ **的条件期望**, 记之为 $\boldsymbol{E}(X|\mathscr{G}) = Y$.

由定义看出, 条件期望 $\boldsymbol{E}(X|\mathscr{G})$ 是 $P$–a.s. 唯一决定的, 即是若有两个 $Y_1$ 和 $Y_2$ 都是 $\mathscr{G}/\mathscr{B}(\mathbf{R})$ 可测的随机变量且

$$Y_i = \boldsymbol{E}(X|\mathscr{G}), \quad (i = 1, 2),$$

则 $Y_1 = Y_2$, $P$–a.s.

特别地,

(1) 若 $\mathscr{G}$ 是实值随机变量族 $\{X_\alpha, \alpha \in \Gamma\}$ 所产生的 $\sigma$ 代数, 即是, $\mathscr{G} = \sigma\left(\bigcup\limits_{\alpha\in\Gamma} X_\alpha^{-1}(\mathscr{B}(\mathbf{R}))\right)$, 记之为 $\mathscr{G} = \sigma(X_\alpha, \alpha \in \Gamma)$, 则定义

$$\boldsymbol{E}(X|X_\alpha, \alpha \in \Gamma) = \boldsymbol{E}(X|\mathscr{G}), \tag{5.2}$$

其中 $\Gamma$ 是任意一个指标集.

(2) 对任一可测集 $A \in \mathscr{F}$ 和 $\mathscr{F}$ 中任一子 $\sigma$ 代数 $\mathscr{G} \subset \mathscr{F}$, 定义 $A$ 关于 $\mathscr{G}$ 的**条件概率**为

$$P(A|\mathscr{G}) = \boldsymbol{E}(\mathbf{1}_A|\mathscr{G}). \tag{5.3}$$

**命题 5.1** 设 $\Omega$ 是任一集合, $(E, \mathscr{E})$ 是任一可测空间, $f: \Omega \mapsto E, g: \Omega \mapsto \overline{\mathbf{R}}$, 则 $g$ 关于 $\sigma(f)$ 可测的充分必要条件是: 存在

$$h: E \mapsto \overline{\mathbf{R}}, \quad h \text{ 关于 } \mathscr{E}/\mathscr{B}(\overline{\mathbf{R}}) \text{ 可测},$$

使 $g = h(f)$.

**证** 充分性显然成立. 下证必要性. 令

$$\mathscr{G} = \{g : g = h(f), \ \ h \text{ 关于 } \mathscr{E}/\mathscr{B}(\overline{\mathbf{R}}) \text{ 可测}\},$$

则 $\mathscr{G}$ 是包含全体常值函数的向量空间. 令 $\{h_n(f), n \geqslant 1\}$ 是 $\mathscr{G}$ 中的使 $F \xlongequal{\text{def.}} \sup\limits_{n\geqslant 1} h_n(f)$ 为实值函数的一个非降序列, 再令

$$A = \{x \in E : \sup_{n\geqslant 1} h_n(x) < \infty\},$$

则 $A \in \mathscr{E}, f(\Omega) \subset A$, 此处 $f(\Omega)$ 如前定义, 为映射 $f$ 在 $\Omega$ 上的像集. 定义

$$h(x) = \begin{cases} \sup\limits_{n\geqslant 1} h_n(x), & \text{当 } x \in A, \\ 0, & \text{反之}, \end{cases}$$

则 $h$ 关于 $\mathscr{E}/\mathscr{B}(\mathbf{R})$ 可测, 且 $F = h(f)$, 即 $F \in \mathscr{G}$. 任取集合 $C \in \sigma(f)$, 必有 $B \in \mathscr{E}$ 使 $C = f^{-1}(B)$, 从而 $\mathbf{1}_C = \mathbf{1}_B(f) \in \mathscr{G}$. 由单调系定理可知 $\mathscr{G}$ 包含了 $\Omega$ 上的一切 $\sigma(f)/\mathscr{B}(\mathbf{R})$ 可测函数, 即是对任何 $\sigma(f)/\mathscr{B}(\mathbf{R})$ 可测的函数 $g$, 必有 $\mathscr{E}/\mathscr{B}(\mathbf{R})$ 可测的函数 $h$, 使 $g = h(f)$.

若 $g$ 是任一 $\sigma(f)/\mathscr{B}(\overline{\mathbf{R}})$ 可测的广义实值函数, 令 $g^* = \arctan g$, 则 $g^*$ 是关于 $\sigma(f)/\mathscr{B}(\mathbf{R})$ 可测的实值函数. 如前所证, 必存在 $\mathscr{E}/\mathscr{B}(\mathbf{R})$ 可测函数 $h^*$, 使 $g^* = h^*(f)$. 可以假定 $h^*(E) \subset \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$. 令 $h = \tan h^*$, 则 $h$ 关于 $\mathscr{E}/\mathscr{B}(\overline{\mathbf{R}})$ 可测, 且 $g = h(f)$. 命题 5.1 证毕.

**命题 5.2** 设 $\Omega$ 为任一集合, $\Gamma$ 为任一指标集 (可以是不可数集), $\{\mathscr{G}_i, i \in \Gamma\}$ 是 $\Omega$ 上的一族 $\sigma$ 代数. 记 $\bigvee\limits_{i\in\Gamma} \mathscr{G}_i = \sigma\left(\bigcup\limits_{i\in\Gamma} \mathscr{G}_i\right)$, 则任取 $A \in \bigvee\limits_{i\in\Gamma} \mathscr{G}_i$, 必存在 $\Gamma$ 的一个可数子集 $I$ (可以依赖 $A$), 使 $A \in \bigvee\limits_{i\in I} \mathscr{G}_i$.

**证** 令

$$\mathfrak{M} = \{A \in \bigvee_{i\in\Gamma} \mathscr{G}_i: \text{存在 } \Gamma \text{ 的可数子集 } I, \text{使 } A \in \bigvee_{i\in I} \mathscr{G}_i\},$$

则 $\mathfrak{M} \supset \bigcup\limits_{i\in\Gamma} \mathscr{G}_i$. 又因为 $\mathfrak{M}$ 是一个 $\sigma$ 代数, 所以 $\mathfrak{M} = \bigvee\limits_{i\in\Gamma} \mathscr{G}_i$. 命题 5.2 证毕.

**命题 5.3** 设 $\Omega$ 和 $\Gamma$ 如前. $\{(E_i, \mathscr{E}_i), i \in \Gamma\}$ 是一族可测空间, $f_i : \Omega \mapsto E_i (i \in \Gamma)$, 令 $\mathscr{F} = \sigma(f_i, i \in \Gamma), \varphi : \Omega \mapsto \overline{\mathbf{R}}$ 且 $\mathscr{F}/\mathscr{B}(\overline{\mathbf{R}})$ 可测, 则存在 $\Gamma$ 的一个可数子集 $I$ (可以依赖于 $\varphi$) 和定义在可测空间 $\left(\mathop{X}\limits_{i\in I} E_i, \mathop{X}\limits_{i\in I} \mathscr{E}_i\right)$ 上的取值于 $\overline{\mathbf{R}}$ 的函数 $h, h$ 关于 $\left(\mathop{X}\limits_{i\in I} \mathscr{E}_i\right)/\mathscr{B}(\overline{\mathbf{R}})$ 可测, 使

$$\varphi = h(f_I),$$

其中 $f_I(\omega) = (f_i(\omega), i \in I), \omega \in \Omega$.

**证** 令 $D$ 为有理数集, 再令

$$\mathfrak{M} = \{[a, b] : a \in D\bigcup\{-\infty\}, b \in D\bigcup\{\infty\}\},$$

则 $\mathfrak{M}$ 为可数个集合构成的集合系, 所以可令

$$\mathfrak{M} = \{A_1, A_2, \cdots\}, \quad A_k \in \mathscr{B}(\overline{\mathbf{R}}).$$

对每个 $A_k$, 由 $\varphi \in \mathscr{F}/\mathscr{B}(\overline{\mathbf{R}})$. 可知 $\varphi^{-1}(A_k) \in \mathscr{F}$. 再用命题 5.2 得知: 存在 $\Gamma$ 的一个可数子集 $I_k$, 使

$$\varphi^{-1}(A_k) \in \sigma(f_i, i \in I_k).$$

令 $I = \bigcup\limits_{k=1}^{\infty} I_k$, 则 $\varphi^{-1}(\mathfrak{M}) \subset \sigma(f_i, i \in I)$, 所以

$$\varphi^{-1}(\mathscr{B}(\overline{\mathbf{R}})) = \varphi^{-1}(\sigma(\mathfrak{M})) = \sigma(\varphi^{-1}(\mathfrak{M})) \subset \sigma(f_i, i \in I),$$

此即

$$\varphi \in \sigma(f_i, i \in I)/\mathscr{B}(\overline{\mathbf{R}}).$$

再利用命题 5.1 得知: 满足命题 5.3 中的条件的 $h$ 是存在的. 命题 5.3 得证.

作为命题 5.3 的应用, 可以得到一个关于条件期望的定理.

**定理 5.1** 设 $Y$ 是完备概率空间 $(\Omega,\mathscr{F},P)$ 上的数学期望存在的实值随机变量, $\Gamma$ 是任一指标集 (可以是不可数集), $\{X_i, i\in\Gamma\}$ 是一族实值随机变量, 则存在 $\Gamma$ 的一个可数子集 $I$ (可以依赖于 $Y$) 及定义在可测空间 $(\underset{i\in I}{\times}\mathbf{R}_i, \underset{i\in I}{\times}\mathscr{B}_i)$ (其中 $\mathscr{B}_i=\mathscr{B}(\mathbf{R}_i), \mathbf{R}_i=\mathbf{R}$) 上的实值函数 $h$, $h\in\left(\underset{i\in I}{\times}\mathscr{B}_i\right)/\mathscr{B}(\mathbf{R})$, 使

$$\boldsymbol{E}(Y|X_i, i\in\Gamma)=h(X_I),$$

其中 $X_I(\omega)=(X_i(\omega), i\in I), \omega\in\Omega$.

**证** 在命题 5.3 中取 $\varphi=\boldsymbol{E}(Y|X_i,i\in\Gamma), E_i=\mathbf{R}, \mathscr{E}_i=\mathscr{B}(\mathbf{R}), f_i=X_i, i\in\Gamma)$, 则由命题 5.3 和命题 5.1 即可得定理 5.1.

**推论 5.1** 若定理 5.1 中的 $\Gamma=\{1,2\cdots,d\}$, 则存在

$$h:\mathbf{R}^d\mapsto\mathbf{R},\quad h\in\mathscr{B}(\mathbf{R}^d)/\mathscr{B}(\mathbf{R}),$$

使

$$\boldsymbol{E}(Y|X_1,\cdots,X_d)=h(X_1,\cdots,X_d).$$

**例 5.1** 设 $X$ 是 $(\Omega,\mathscr{F},P)$ 上的实值随机变量, 且 $\boldsymbol{E}(X)$ 存在, $\mathscr{G}=\{\varnothing,\Omega\}$, 则

$$\boldsymbol{E}(X|\mathscr{G})=\boldsymbol{E}(X).$$

**例 5.2** 设 $X$ 和 $Y$ 都是 $(\Omega,\mathscr{F},P)$ 上取非负整数的随机变量, 其联合分布为 $P(X=i,Y=j)=p_{i,j}, Y$ 的边缘分布为 $P(Y=j)=p_j>0, i,j\in\mathbf{Z}_+$. 则存在 Borel 可测函数 $f$, 使

$$\boldsymbol{E}(X|Y)=f(Y)\quad P-\text{a.s.},$$

且 $f(j)=\left(\sum\limits_{i=0}^{\infty}ip_{i,j}\right)/p_j, (j\in\mathbf{Z}_+)$.

注意: 此时 $f(j)=\sum\limits_{i=0}^{\infty}P(X=i|Y=j)i$, 这与初等概率论中定义的条件期望是一致的.

**例 5.3** 设 $X,Y$ 均为 $(\Omega,\mathscr{F},P)$ 上的实值随机变量, $(X,Y)$ 有联合密度函数 $f(x,y)$ 即是对任何 $A\in\mathscr{B}(\mathbf{R}^2)$, 有

$$P((X,Y)\in A)=\int_A f(x,y)dxdy.$$

令

$$f_2(y)=\int_{\mathbf{R}}f(x,y)dx,$$
$$C=\{y\in\mathbf{R}:f_2(y)>0\},$$
$$g(x,y)=\begin{cases}f(x,y)/f_2(y), & \text{当 } x\in\mathbf{R},y\in C,\\ 0, & \text{当 } x\in\mathbf{R},y\overline{\in}C.\end{cases}$$

则

$$\boldsymbol{E}(X|Y)=\int_{\mathbf{R}}xg(x,Y)dx\quad P\text{-a.s.},$$

这与初等概率论中定义的条件期望也是一致的.

由于条件期望 (概率) 都是差一个 $P$ 零测集唯一决定的, 所以以后谈及条件期望 (概率) 相等时, 均系除一个 $P$ 零测集以外相等, 故有时 $P$-a.s. 符号略去不写.

下面我们介绍条件期望 (概率) 的一些简单性质:

(1) 设 $X,X_1,\cdots,X_n$ 是 $(\Omega,\mathscr{F},P)$ 上的实值随机变量, $c,c_1,\cdots,c_n$ 是实数, $\mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, $\boldsymbol{E}(|X|)<\infty,\boldsymbol{E}(|X_i|)<\infty,i=1,\cdots,n$. 则

(a) $\boldsymbol{E}(c|\mathscr{G})=c$;

(b) $\boldsymbol{E}\left(\sum\limits_{i=1}^{n}c_iX_i|\mathscr{G}\right)=\sum\limits_{i=1}^{n}c_i\boldsymbol{E}(X_i|\mathscr{G})$;

(c) $X_1\geqslant X_2\Rightarrow\boldsymbol{E}(X_1|\mathscr{G})\geqslant\boldsymbol{E}(X_2|\mathscr{G})$.

(2) 设 $X$ 如 (1), $\mathscr{G}_1,\mathscr{G}_2,\mathscr{G}_3$ 皆为 $\mathscr{F}$ 的子 $\sigma$ 代数, 且 $\mathscr{G}_1\subset\mathscr{G}_2\subset\mathscr{G}_3$, 如果 $\boldsymbol{E}(X|\mathscr{G}_3)$ 关于 $\mathscr{G}_1/\mathscr{B}(\mathbf{R})$ 可测, 则

$$\boldsymbol{E}(X|\mathscr{G}_1)=\boldsymbol{E}(X|\mathscr{G}_2)=\boldsymbol{E}(X|\mathscr{G}_3),$$
$$\boldsymbol{E}(\boldsymbol{E}(X|\mathscr{G}_2)|\mathscr{G}_1)=\boldsymbol{E}(\boldsymbol{E}(X|\mathscr{G}_1)|\mathscr{G}_2)=\boldsymbol{E}(X|\mathscr{G}_1).$$

(3) 设 $\mathscr{G}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 中 $\mathscr{F}$ 的子 $\sigma$ 代数, $A_n\in\mathscr{F},n=1,2,\cdots$, 则

(a) $A_n\uparrow A\Rightarrow P(A_n|\mathscr{G})\uparrow P(A|\mathscr{G})$;

$$A_n\downarrow A\Rightarrow P(A_n|\mathscr{G})\downarrow P(A|\mathscr{G});$$

(b) $\{A_n\}$ 两两不交 $\Rightarrow\sum\limits_{n=1}^{\infty}P(A_n|\mathscr{G})=P\left(\bigcup\limits_{n=1}^{\infty}A_n|\mathscr{G}\right)$;

(4) 设 $X,Y$ 都是 $(\Omega,\mathscr{F},P)$ 上的实值随机变量, $\boldsymbol{E}(|X|)<\infty,\boldsymbol{E}(|XY|)<\infty$, $\mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, $Y$ 关于 $\mathscr{G}/\mathscr{B}(\mathbf{R})$ 可测, 则

$$\boldsymbol{E}(XY|\mathscr{G})=Y\boldsymbol{E}(X|\mathscr{G}).$$

特别地, 若 $f$ 是实变实值的 Borel 可测函数, 则

$$\boldsymbol{E}(Xf(Y)|Y) = f(Y)\boldsymbol{E}(X|Y). \tag{5.4}$$

(5) 若 $X$ 是 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量, $\boldsymbol{E}(|X|) < \infty, \mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, $\mathscr{G}$ 与 $\sigma(X)$ 相互独立, 即是对任何 $A \in \mathscr{G}, B \in \sigma(X)$, 有 $P(A \bigcap B) = P(A)P(B)$, 则

$$\boldsymbol{E}(X|\mathscr{G}) = \boldsymbol{E}(X). \tag{5.5}$$

(6) 设 $X$ 是 $(\Omega, \mathscr{F}, P)$ 上的实值随机变量, $\boldsymbol{E}(|X|) < \infty, \mathscr{G}$ 是 $\mathscr{F}$ 中的 $\sigma$ 代数, $A \in \mathscr{G}, \mathscr{G}_1 \stackrel{\text{def.}}{=\!=} \{B : B = AD, D \in \mathscr{G}\}$, 则 $\mathscr{G}_1$ 是 $\mathscr{G}$ 中的子 $\sigma$ 代数, $A \in \mathscr{G}_1$, 且

$$\boldsymbol{E}(X|\mathscr{G})\mathbf{1}_A = \boldsymbol{E}(X|\mathscr{G}_1)\mathbf{1}_A. \tag{5.6}$$

用条件期望之定义立即可验证 (1)~(5) 成立. 下面只证 (5.6) 式.

由于 (5.6) 两边都关于 $\mathscr{G}/\mathscr{B}(\mathbf{R})$ 可测, 又因为任取 $D \in \mathscr{G}, AD \in \mathscr{G}_1$, 所以

$$\begin{aligned} &\int_D \boldsymbol{E}(X|\mathscr{G})\mathbf{1}_A dP = \int_D \boldsymbol{E}(\mathbf{1}_A X|\mathscr{G})dP \\ &= \int_D \mathbf{1}_A X dP = \int_{AD} X dP = \int_{AD} \boldsymbol{E}(X|\mathscr{G}_1)dP \\ &= \int_D \boldsymbol{E}(X|\mathscr{G}_1)\mathbf{1}_A dP, \end{aligned}$$

所以 (5.6) 成立. 性质 (6) 证毕.

**定义 5.2** 设 $(\Omega, \mathscr{F}, P)$ 为完备概率空间, $\mathscr{G}, \mathscr{G}_1, \mathscr{G}_2$ 都是 $\mathscr{F}$ 中的子 $\sigma$ 代数. 若对任何 $A_1 \in \mathscr{G}_1, A_2 \in \mathscr{G}_2$, 都有

$$P(A_1A_2) = P(A_1)P(A_2), \tag{5.7}$$

则称 $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 相互独立; 若对任何 $A_1 \in \mathscr{G}_1, A_2 \in \mathscr{G}_2$, 都有

$$P(A_1|\mathscr{G})P(A_2|\mathscr{G}) = P(A_1A_2|\mathscr{G}), \tag{5.8}$$

则称 $\mathscr{G}_1$ **与** $\mathscr{G}_2$ **关于** $\mathscr{G}$ **条件独立**.

对于任一抽象空间 $\Omega$ 中的一族 $\sigma$ 代数, $\{\mathscr{G}_\alpha, \alpha \in \Gamma\}$, $\bigcup\limits_{\alpha\in\Gamma} \mathscr{G}_\alpha$ 未必是 $\sigma$ 代数, 但 $\sigma\left(\bigcup\limits_{\alpha\in\Gamma} \mathscr{G}_\alpha\right)$ 是包含一切 $\mathscr{G}_\alpha(\alpha \in \Gamma)$ 的最小 $\sigma$ 代数. 本书恒用 $\bigvee\limits_{\alpha\in\Gamma} \mathscr{G}_\alpha$ 表 $\sigma\left(\bigcup\limits_{\alpha\in\Gamma} \mathscr{G}_\alpha\right)$.

**定理 5.2** 设 $\mathscr{G},\mathscr{G}_1$ 和 $\mathscr{G}_2$ 之定义如定义 5.2, 则 $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 关于 $\mathscr{G}$ 条件独立的充分必要条件是

$$P(B_2|\mathscr{G}) = P(B_2|\mathscr{G} \vee \mathscr{G}_1), \quad (\forall B_2 \in \mathscr{G}_2). \tag{5.9}$$

**证** 任取 $B \in \mathscr{G}, B_i \in \mathscr{G}_i (i=1,2)$. 由条件期望的定义和条件期望的性质 (4) 有:

$$\int_{BB_1} \mathbf{1}_{B_2} dP = \int_B \mathbf{1}_{B_1 B_2} dP = \int_B \boldsymbol{E}(\mathbf{1}_{B_1}\mathbf{1}_{B_2}|\mathscr{G}) dP, \tag{5.10}$$

$$\begin{aligned}\int_{BB_1} P(B_2|\mathscr{G}) dP &= \int_B \boldsymbol{E}\left(\mathbf{1}_{B_1}\boldsymbol{E}(\mathbf{1}_{B_2}|\mathscr{G})|\mathscr{G}\right) dP \\ &= \int_B \boldsymbol{E}(\mathbf{1}_{B_1}|\mathscr{G})\boldsymbol{E}(\mathbf{1}_{B_2}|\mathscr{G}) dP.\end{aligned} \tag{5.11}$$

首先证明定理 5.2 的必要性部分. 设 $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 关于 $\mathscr{G}$ 条件独立, 则对一切 $B \in \mathscr{G}, B_i \in \mathscr{G}_i, (i=1,2,)$, 恒有

$$\boldsymbol{E}(\mathbf{1}_{B_1}|\mathscr{G})\boldsymbol{E}(\mathbf{1}_{B_2}|\mathscr{G}) = \boldsymbol{E}(\mathbf{1}_{B_1}\mathbf{1}_{B_2}|\mathscr{G}). \tag{5.12}$$

由 (5.10)—(5.12) 有

$$\int_{BB_1} \mathbf{1}_{B_2} dP = \int_{BB_1} P(B_2|\mathscr{G}) dP. \tag{5.13}$$

欲证 (5.9) 成立, 注意 $P(B_2|\mathscr{G})$ 关于 $\mathscr{G} \vee \mathscr{G}_1$ 可测, 则只需证明对任何 $C \in \mathscr{G} \vee \mathscr{G}_1$, 总有

$$\int_C \mathbf{1}_{B_2} dP = \int_C P(B_2|\mathscr{G}) dP. \tag{5.14}$$

事实上, 对任何 $B_2 \in \mathscr{G}_2$, 令

$$\mathfrak{M} = \{C : \int_C \mathbf{1}_{B_2} dP = \int_C P(B_2|\mathscr{G}) dP\},$$
$$\mathfrak{M}_1 = \{C : C = BB_1, B \in \mathscr{G}, B_1 \in \mathscr{G}_1\},$$

则 $\mathfrak{M}_1$ 是 $\pi$ 系且 $\mathfrak{M}_1 \supset \mathscr{G} \bigcup \mathscr{G}_1$. 由 (5.13) 知 $\mathfrak{M} \supset \mathfrak{M}_1$. 显然 $\mathfrak{M}$ 是 $d$ 系, 所以由第一章定理 1.2 知

$$\mathfrak{M} \supset d(\mathfrak{M}_1) = \sigma(\mathfrak{M}_1) \supset \sigma(\mathscr{G} \bigcup \mathscr{G}_1) = \mathscr{G} \vee \mathscr{G}_1.$$

此即 (5.14) 对一切 $C \in \mathscr{G} \vee \mathscr{G}_1$ 成立. 必要性得证.

最后证明定理 5.2 的充分性部分. 设 (5.9) 对一切 $B_2 \in \mathscr{G}_2$ 成立, 则

$$\int_{BB_1} P(B_2|\mathscr{G}) dP = \int_{BB_1} \mathbf{1}_{B_2} dP, \quad (B \in \mathscr{G}, B_i \in \mathscr{G}_i). \tag{5.15}$$

由 (5.10)、(5.11) 和 (5.15) 有

$$\int_B \boldsymbol{E}(\mathbf{1}_{B_1}\mathbf{1}_{B_2}|\mathscr{G})dP$$
$$= \int_B \boldsymbol{E}(\mathbf{1}_{B_1}|\mathscr{G})\boldsymbol{E}(\mathbf{1}_{B_2}|\mathscr{G})dP, \quad (B \in \mathscr{G}, B_i \in \mathscr{G}_i). \tag{5.16}$$

而 (5.16) 左右两边的被积函数关于 $\mathscr{G}$ 都是可测的, 所以

$$\boldsymbol{E}(\mathbf{1}_{B_1}\mathbf{1}_{B_2}|\mathscr{G}) = \boldsymbol{E}(\mathbf{1}_{B_1}|\mathscr{G})\boldsymbol{E}(\mathbf{1}_{B_2}|\mathscr{G}).$$

充分性得证. 定理证毕.

下面我们给出期望和条件期望的一个重要的不等式 — Jensen 不等式. 首先给出凸集与凸函数的定义和几个简单性质.

**定义 5.3** 称 $\mathbf{R}^d$ 中的子集 $S$ 是凸集, 如果:

$$x, y \in S \Rightarrow ((1-\lambda)x + \lambda y) \in S, \quad (\forall \lambda \in (0,1)).$$

设 $A$ 是 $\mathbf{R}^d$ 中的一个子集, $f: A \mapsto \mathbf{R}$, 若

$$\{(x,t): x \in A, x \in \mathbf{R}, t \geqslant f(x)\}$$

是 $\mathbf{R}^{d+1}$ 中的**凸集**, 则称 $f$ 是 $A$ 上的**凸函数**.

**命题 5.1** $\mathbf{R}$ 中的子集 $S$ 是凸集的充分必要条件是: $S$ 是 $\mathbf{R}$ 中的区间 (开的或闭的或半开半闭的, 有穷的或无穷的).

**命题 5.2** 设 $f: (\alpha,\beta) \mapsto \mathbf{R}$, ($\alpha$ 可为 $-\infty$, $\beta$ 可为 $\infty$), 则

(1) $f$ 是 $(\alpha,\beta)$ 上的凸函数的充分必要条件是: $f$ 在 $(\alpha,\beta)$ 上连续而且

$$f\left(\frac{x+y}{2}\right) \leqslant \frac{f(x)+f(y)}{2}, \quad (\forall x, y \in (\alpha,\beta)),$$

从而 $(\alpha,\beta)$ 上的凸函数 $f$ 必是 Borel 可测的, 即是, $f \in ((\alpha,\beta)\bigcap\mathscr{B}(\mathbf{R}))/\mathscr{B}(\mathbf{R})$;

(2) $f$ 在 $(\alpha,\beta)$ 上有单调非降连续导函数 $f' \Rightarrow f$ 是 $(\alpha,\beta)$ 上的凸函数;

(3) $f(x) = |x|^p, (1 \leqslant p < \infty) \Rightarrow f$ 是 $\mathbf{R}$ 上的凸函数.

**命题 5.3** 设 $f$ 是 $\mathbf{R}$ 中的凸子集 $S$ 上的凸函数, 则必存在线性函数列 $\{f_n(x) = a_n x + b_n, n \geqslant 1\}$, 使 $f_n(x) \leqslant f(x), (\forall n \geqslant 1, x \in S), f(x) = \sup\limits_{n\geqslant 1} f_n(x)$. ($\forall x \in S^\circ$, $S^\circ$ 是 $S$ 的开核, 即含于 $S$ 的最大开集.)

命题 5.1~5.3 的证明请参见 [102], 在此从略.

**定理 5.3 (Jensen 不等式)** 设 $X$ 是完备概率空间 $(\Omega, \mathscr{F}, P)$ 的实值随机变量, $f$ 是 $\mathbf{R}$ 中的凸子集 $S$ 上的凸函数, $X(\Omega) \subset S, \boldsymbol{E}(X) < \infty, \mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, 则

(1) $f(\boldsymbol{E}(X)) \leqslant \boldsymbol{E}(f(X))$;

(2) $f(\boldsymbol{E}(X|\mathscr{G})) \leqslant \boldsymbol{E}(f(X)|\mathscr{G})$.

**证** 由命题 5.1, 可令 $S^\circ = (\alpha, \beta), S - S^\circ$ 至多只含两点 $\alpha$ 和 $\beta$ (当 $\alpha, \beta$ 为有限实数才有可能). 而由 $X(\Omega) \subset S$ 可知 $\boldsymbol{E}(X) \in S$.

先证结论 (1).

(A) 若 $\boldsymbol{E}(X) \in S^\circ$, 则取线性函数列 $\{f_n\}$ 满足命题 5.3 中的条件. 于是有:

$$\begin{aligned} f(\boldsymbol{E}(X)) &= \sup_{n\geqslant 1} f_n(\boldsymbol{E}(X)) = \sup_{n\geqslant 1} \boldsymbol{E}(f_n(X)) \\ &\leqslant \boldsymbol{E}(f(X)). \end{aligned}$$

(B) 若 $\boldsymbol{E}(X) \in S - S^\circ$, 不妨设 $\boldsymbol{E}(X) = \alpha$ 是实数. 由于 $X(\omega) \in [\alpha, \beta] (\forall \omega \in \Omega)$, 所以 $X = \alpha$, a.s., 从而 $f(\boldsymbol{E}(X)) = f(\alpha) = \boldsymbol{E}(f(X))$.

再证结论 (2).

取线性函数列 $\{f_n\}$ 满足命题 5.3 中的条件, 并令

$$\begin{aligned} A^\circ &= \{\omega \in \Omega : \boldsymbol{E}(X|\mathscr{G})(\omega) \in S^\circ\}; \\ A_\alpha &= \{\omega \in \Omega : \boldsymbol{E}(X|\mathscr{G})(\omega) = \alpha\}; \\ A_\beta &= \{\omega \in \Omega : \boldsymbol{E}(X|\mathscr{G})(\omega) = \beta\}. \end{aligned}$$

(A) 若 $\omega \in A^\circ$, 则

$$\begin{aligned} f(\boldsymbol{E}(X|\mathscr{G})(\omega)) &= \sup_{n\geqslant 1} f_n(\boldsymbol{E}(X|\mathscr{G})(\omega)) \\ &= \sup_{n\geqslant 1} \boldsymbol{E}(f_n(X)|\mathscr{G})(\omega) \\ &\leqslant \boldsymbol{E}(f(X)|\mathscr{G})(\omega). \end{aligned}$$

(B) 若 $\omega \in A_\alpha$, 则由 $X - \alpha \geqslant 0, A_\alpha \in \mathscr{F}$ 有

$$\int_{A_\alpha} (X - \alpha) dP = \int_{A_\alpha} (\boldsymbol{E}(X|\mathscr{G}) - \alpha) dP = 0,$$

所以 $X = \alpha, \text{a.s.} in A_\alpha$. 而由 $A_\alpha$ 的定义还有

$$\begin{aligned} \mathbf{1}_{A_\alpha} f(\boldsymbol{E}(X|\mathscr{G})) &= \mathbf{1}_{A_\alpha} f(\alpha) = \boldsymbol{E}(\mathbf{1}_{A_\alpha} f(\alpha)|\mathscr{G}) \\ &= \boldsymbol{E}(\mathbf{1}_{A_\alpha} f(X)|\mathscr{G}) = \mathbf{1}_{A_\alpha} \boldsymbol{E}(f(X)|\mathscr{G}), \end{aligned}$$

此即

$$f(\boldsymbol{E}(X|\mathscr{G})(\omega))=\boldsymbol{E}(f(X)|\mathscr{G})(\omega),\quad (\forall\omega\in A_\alpha).$$

(C) 仿 (B) 可证

$$f(\boldsymbol{E}(X|\mathscr{G})(\omega))=\boldsymbol{E}(f(X)|\mathscr{G})(\omega),\quad (\forall\omega\in A_\beta).$$

定理 5.3 证毕.

## *§6 随机元的条件期望

在这一节中, 我们将要研究取值于 Banach 空间的随机元的条件期望.

**定义 6.1** ($\boldsymbol{B}$ 值强可测随机元的条件期望) 设 $(\Omega,\mathscr{F},P)$ 是完备概率空间,$(\boldsymbol{B},\|\bullet\|)$ 是 Banach 空间, $X$ 是 $\boldsymbol{B}$ 值强可测随机元且 Bochner 可积, $\mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数而且也是完备的 (即 $\mathscr{G}$ 中的 $P$ 零测集之子集亦属于 $\mathscr{G}$). 如果存在

$$Y:\Omega\mapsto\boldsymbol{B}$$

满足以下诸条件:

- (1) $Y$ 关于 $\mathscr{G}$ 强可测且 $\boldsymbol{E}(\|Y\|)<\infty$ (从而 $Y$ 是 Bochner 可积的);

(2) $(\mathrm{s})\int_A XdP=(\mathrm{s})\int_A YdP\quad(\forall A\in\mathscr{G})$,

则称 $Y$ 是 $X$ 关于 $\sigma$ 代数 $\mathscr{G}$ 的**强条件期望** (简称条件期望), 记作

$$(\mathrm{s})\boldsymbol{E}(X|\mathscr{G})=Y.$$

若 $\{X_i,i\in\Gamma\}$ 是一族 $\boldsymbol{B}$ 值随机元, 则记

$$(\mathrm{s})\boldsymbol{E}(X|X_i,i\in\Gamma)=(\mathrm{s})\boldsymbol{E}(X|\sigma(X_i,i\in\Gamma)).$$

**定理 6.1** 设 $(\Omega,\mathscr{F},P),(\boldsymbol{B},\|\bullet\|)$, $X$ 和 $\mathscr{G}$ 如定义 6.1 所定义, 则 $(\mathrm{s})\boldsymbol{E}(X|\mathscr{G})$ 在 $P$–a.s. 意义下是唯一存在的.

**证** 若 $X=\sum\limits_{i=1}^n x_i\mathbf{1}_{A_i}$ 是简单的 $\boldsymbol{B}$ 值函数, 令

$$Y=\sum_{i=1}^n x_i\boldsymbol{E}(\mathbf{1}_{A_i}|\mathscr{G}),$$

则 $Y$ 关于 $\mathscr{G}$ 强可测, $\boldsymbol{E}(\|Y\|)<\infty$, 且对任何 $A\in\mathscr{G}$, 总有

$$\begin{aligned}(\mathrm{s})\int_A YdP&=\sum_{i=1}^n x_i\int_A\boldsymbol{E}(\mathbf{1}_{A_i}|\mathscr{G})dP\\&=\sum_{i=1}^n x_i\int_A\mathbf{1}_{A_i}dP=(\mathrm{s})\int_A XdP,\end{aligned}\tag{6.1}$$

此即 $Y=(\mathrm{s})\boldsymbol{E}(X|\mathscr{G})$.

若 $X$ 是一般的 Bochner 可积的强可测的 $\boldsymbol{B}$ 值随机元, 则由命题 4.4 得知: 存在简单 $\boldsymbol{B}$ 值函数列

$$X_n=\sum_{i=1}^{k_n}x_i^{(n)}\mathbf{1}_{A_i^{(n)}}\quad(n\geqslant 1),$$

使

$$\|X_n(\omega)\|\leqslant 2\|X(\omega)\|,\quad(\forall\omega\in\Omega,n\geqslant 1),\tag{6.2}$$

$$\lim_{n\to\infty}X_n=X,\quad\text{a.s.},\tag{6.3}$$

$$\lim_{n\to\infty}\boldsymbol{E}(\|X-X_n\|)=0.\tag{6.4}$$

令

$$Y_n=(\mathrm{s})\boldsymbol{E}(X_n|\mathscr{G})=\sum_{i=1}^{k_n}x_i^{(n)}\boldsymbol{E}(\mathbf{1}_{A_i^{(n)}}|\mathscr{G}),$$

则

$$Y_n\text{ 关于 }\mathscr{G}\text{ 强可测},\quad\boldsymbol{E}(\|Y_n\|)<\infty.$$

又因为

$$X_n-X_m=\sum_{i=1}^{k_n}\sum_{j=1}^{k_m}(x_i^{(n)}-x_j^{(m)})\mathbf{1}_{A_i^{(n)}}\mathbf{1}_{A_j^{(m)}},$$

$$Y_n-Y_m=\sum_{i=1}^{k_n}\sum_{j=1}^{k_m}(x_i^{(n)}-x_j^{(m)})\boldsymbol{E}(\mathbf{1}_{A_i^{(n)}}\mathbf{1}_{A_j^{(m)}}|\mathscr{G}),$$

所以

$$\begin{aligned}\boldsymbol{E}(\|Y_n-Y_m\|)&\leqslant\sum_{i=1}^{k_n}\sum_{j=1}^{k_m}\|x_i^{(n)}-x_j^{(m)}\|P(A_i^{(n)}A_j^{(m)})\\&=\boldsymbol{E}(\|X_n-X_m\|).\end{aligned}\tag{6.5}$$

因此

$$\lim_{n,m\to\infty}\boldsymbol{E}(\|Y_n-Y_m\|)=0.\tag{6.6}$$

而 $L^1(\Omega,\mathscr{F},P;\boldsymbol{B})\overset{\text{def.}}{=\!=}\{X:\Omega\mapsto\boldsymbol{B},\ X\text{ 是强可测的且 }\boldsymbol{E}(\|X\|)<\infty\}$ 是完备的, 所以必存在 $Y\in L^1(\Omega,\mathscr{F},P;\boldsymbol{B})$, 使

$$\lim_{n\to\infty}\boldsymbol{E}(\|Y_n-Y\|)=0.\tag{6.7}$$

下面证明

$$Y = (\mathrm{s})\boldsymbol{E}(X|\mathscr{G}). \tag{6.8}$$

由于 $Y \in L'(\Omega, \mathscr{F}, P; \boldsymbol{B})$, 所以为证 (6.8), 只需证明

$$(\mathrm{s})\int_A XdP = (\mathrm{s})\int_A YdP, \quad (\forall A \in \mathscr{G}). \tag{6.9}$$

但是

$$\lim_{n\to\infty} \boldsymbol{E}(\|X - X_n\|) = \lim_{n\to\infty} \boldsymbol{E}(\|Y - Y_n\|) = 0,$$

所以

$$\begin{aligned}&\lim_{n\to\infty}\int_A \|X - X_n\|dP\\ =&\lim_{n\to\infty}\int_A \|Y - Y_n\|dP = 0 \quad (\forall A \in \mathscr{G}).\end{aligned} \tag{6.10}$$

因此, 由 $Y_n = (\mathrm{s})\boldsymbol{E}(X_n|\mathscr{G})$ 可得

$$\begin{aligned}&(\mathrm{s})\int_A YdP = \lim_{n\to\infty}\left((\mathrm{s})\int_A Y_n dP\right)\\ =&\lim_{n\to\infty}\left((\mathrm{s})\int_A X_n dP\right)\\ =&(\mathrm{s})\int_A XdP \quad (\forall A \in \mathscr{G}).\end{aligned}$$

(6.9) 式得证.

最后证明条件期望在 $P$-a.s. 的意义下唯一决定. 为此, 证明下面的

**命题 6.1** 设 $(\Omega, \mathscr{F}, P), (\boldsymbol{B}, \|\bullet\|), \mathscr{G}$ 如定理 6.1 所定义, $Y$ 是 Bochner 可积的 $\boldsymbol{B}$ 值随机元, 而且关于 $\mathscr{G}$ 强可测. 若 $(\mathrm{s})\int_A YdP = 0 (\forall A \in \mathscr{G})$, 则 $Y = 0$, $P$-a.s..

**证** 由命题 4.4 得知: 可取

$$\begin{aligned}&Y_n = \sum_{i=1}^{k_n} y_i^{(n)} \mathbf{1}_{A_i^{(n)}} \quad (n \geqslant 1),\\ &A_i^{(n)} \in \mathscr{G}, A_i^{(n)} A_j^{(n)} = \varnothing \quad (i \neq j, n \geqslant 1),\\ &\bigcup_{i=1}^{k_n} A_i^{(n)} = \Omega, \{y_i^{(n)} : i = 1, \cdots, k_n, n \geqslant 1\} \subset \boldsymbol{B},\\ &\|Y_n(\omega)\| \leqslant 2\|Y(\omega)\| \quad (\forall \omega \in \Omega, n \geqslant 1),\\ &\lim_{n\to\infty} Y_n = Y, \quad P-\text{a.s.},\\ &\lim_{n\to\infty} \boldsymbol{E}(\|Y_n - Y\|) = 0.\end{aligned}$$

由 $A_i^{(n)} \in \mathscr{G}$ 及命题假设可知:

$$\begin{aligned}
\int_\Omega \|Y_n\| dP &= \sum_{i=1}^{k_n} \int_{A_i^{(n)}} \|Y_n\| dP \\
&= \sum_{i=1}^{k_n} \|y_i^{(n)}\| P(A_i^{(n)}) \\
&= \sum_{i=1}^{k_n} \left\| (\mathrm{s}) \int_{A_i^{(n)}} Y_n dP \right\| \\
&\leqslant \sum_{i=1}^{k_n} \left\| (\mathrm{s}) \int_{A_i^{(n)}} Y_n dP - (\mathrm{s}) \int_{A_i^{(n)}} Y dP \right\| \\
&\leqslant \sum_{i=1}^{k_n} \int_{A_i^{(n)}} \|Y_n - Y\| dP \\
&= \int_\Omega \|Y_n - Y\| dP.
\end{aligned}$$

所以

$$\begin{aligned}
\int_\Omega \|Y\| dP &\leqslant \int_\Omega \|Y - Y_n\| dP + \int_\Omega \|Y_n\| dP \\
&\leqslant 2 \int_\Omega \|Y - Y_n\| dP \to 0 \quad (\text{当 } n \to \infty \text{ 时}).
\end{aligned}$$

故 $Y = 0$, $P$-a.s.. 命题 6.1 证毕, 从而定理 6.1 得证.

下面介绍强条件期望的性质, 证明甚易, 读者可作为习题验证之.

**命题 6.2** 设 $(\Omega, \mathscr{F}, P)$ 是完备概率空间, $(\boldsymbol{B}, \|\cdot\|)$ 是 Banach 空间, $X, X_1, \cdots, X_n, \cdots$ 是 Bochner 可积的强可测 $\boldsymbol{B}$ 值随机元, $\mathscr{G}, \mathscr{G}_1, \mathscr{G}_2, \mathscr{G}_3$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, $\alpha_1, \cdots, \alpha_n$ 是实数, $x \in \boldsymbol{B}, Y$ 是强可测 $\boldsymbol{B}$ 值随机元, 则

(1) $(\mathrm{s})\boldsymbol{E}(x|\mathscr{G}) = x$;

(2) $(\mathrm{s})\boldsymbol{E}\left(\sum\limits_{i=1}^n \alpha_i X_i|\mathscr{G}\right) = \sum\limits_{i=1}^n \alpha_i (\mathrm{s})\boldsymbol{E}(X_i|\mathscr{G})$;

(3) $\|(\mathrm{s})\boldsymbol{E}(X|\mathscr{G})\| \leqslant \boldsymbol{E}(\|X\||\mathscr{G})$;

(4) 若 $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{G}_3$, 而且 $(\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_3)$ 关于 $\mathscr{G}_1$ 强可测, 则

$$\begin{aligned}
(\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_1) &= (\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_2) = (\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_3); \\
&= (\mathrm{s})\boldsymbol{E}(((\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_2))|\mathscr{G}_1) \\
&= (\mathrm{s})\boldsymbol{E}(((\mathrm{s})\boldsymbol{E}(X|\mathscr{G}_1))|\mathscr{G}_2);
\end{aligned}$$

(5) 若 $X$ 与 $Y$ 相互独立, 则

$$(\mathrm{s})\boldsymbol{E}(X|Y) = (\mathrm{s})\boldsymbol{E}(X);$$

(6) 若 $\|X_n\| \leqslant \|X\|, (\forall n \geqslant 1)$, $\lim\limits_{n\to\infty} X_n = X$, $P$–a.s., 则

$$\lim_{n\to\infty} (\mathrm{s})\boldsymbol{E}(X_n|\mathscr{G}) = (\mathrm{s})\boldsymbol{E}(X|\mathscr{G}).$$

## §7 习题及应用

1. 设 $a_i$ 为实数, $r_i$ 为有理数, $i = 1, \cdots, N$,

$\mathfrak{M}_1 = \{\mathbf{R}^N$ 中一切区间 $(a, b]\}$;

$\mathfrak{M}_2 = \{\mathbf{R}^N$ 中一切闭区间 $[a, b]\}$;

$\mathfrak{M}_3 = \{\mathbf{R}^N$ 中一切开区间 $(a, b)\}$;

$\mathfrak{M}_4 = \{\mathbf{R}^N$ 中一切开集$\}$;

$\mathfrak{M}_5 = \{\mathbf{R}^N$ 中一切闭集 $\}$;

$\mathfrak{M}_6 = \{\mathbf{R}^N$ 中一切形如 $\{x \in \mathbf{R}^N : x_i < a_i, i = 1, \cdots, N\}$ 之集合$\}$;

$\mathfrak{M}_7 = \{\mathbf{R}^N$ 中一切形如 $\{X \in \mathbf{R}^N : x_i \leqslant a_i, i = 1, \cdots N\}$ 之集合$\}$;

$\mathfrak{M}_8 = \{\mathbf{R}^N$ 中一切形如 $\{x \in \mathbf{R}^N : x_i < r_i, i = 1, \cdots, N\}$ 之集合$\}$;

$\mathfrak{M}_9 = \{\mathbf{R}^N$ 中一切形如 $\{x \in \mathbf{R}^N : x_i \leqslant r_i, i = 1, \cdots, N\}$ 之集合$\}$,

则

$$\mathscr{B}(\mathbf{R}^N) = \sigma(\mathfrak{M}_i), i = 1, \cdots, 9.$$

2. 设 $X : \Omega_1 \to \Omega_2, \mathfrak{M}$ 是 $\Omega_2$ 的一个子集族. 记 $X^{-1}(\mathfrak{M}) = \{A : A = X^{-1}(B), B \in \mathfrak{M}\}$ 是 $\Omega_1$ 中的 $X$ 在 $\mathfrak{M}$ 上的逆像子集族, 则

$$\sigma(X^{-1}(\mathfrak{M})) = X^{-1}(\sigma(\mathfrak{M})).$$

3. 设 $(\Omega, \mathscr{F}, P)$ 是概率空间, $(\boldsymbol{B}, \mathscr{B})$ 是任一可测空间, $X : \Omega \to \boldsymbol{B}$, 则 $X$ 是 $\boldsymbol{B}$ 值随机元的充分必要条件是: $X^{-1}(\mathfrak{M}) \subset \mathscr{F}$, 且 $\sigma(\mathfrak{M}) = \mathscr{B}$.

4. 设 $F(x)$ 是分布函数, $k \neq 0$, 则函数

$$\Phi(x) = \frac{1}{k}\int_x^{x+k} F(y)dy, \quad \Psi(x) = \frac{1}{2k}\int_{x-k}^{x+k} F(y)dy$$

都是分布函数.

5. 通过对直径的近似测量来估计圆面积的分布. 设直径的测量值均匀地分布于 $(a, b)$ 之间. 试求圆面积的数学期望.

6. 假定有 $N$ 个人去验血 ($N$ 相当大). 于是有两种方法可以采用: (i) 每个人单独地去检验一次, 于是共需检验 $N$ 次; (ii) 把 $N$ 个人分成 $K$ 人一组, 把这

$K$ 个人的血混在一起去化验. 如果化验结果是阴性反应, 则对这 $K$ 个人只需化验一次就可下结论了. 如果是阳性反应, 则这 $K$ 个人再逐个地化验一次, 这样, 一共要化验 $K+1$ 次才能对这 $K$ 个人下结论.

如果每个人化验结果为阳性的概率为 $p$, 试求:

(a) 用方案 (ii), 所需化验的次数 $X_K$ 的数学期望 $\boldsymbol{E}(X_K)$;

(b) $K$ 取何值方能使 $\boldsymbol{E}(X_K)$ 达最小?

7. 设 $Y$ 和 $X_1, X_2, \cdots$ 为相互独立的取非负整数的随机变量序列, $\{X_n, n \geqslant 0\}$ 具有公共分布, 其公共的母函数为 $f(s), Y$ 的母函数为 $g(s)$. 令

$$X = \begin{cases} X_1 + X_2 + \cdots + X_Y, & \text{当 } Y > 0, \\ 0, & \text{当 } Y = 0. \end{cases}$$

试求 $X$ 的母函数.

8. 设 $\{a_i, i \geqslant 0\}$ 是一列非负的界于 1 的实数, 而且对每个 $n \geqslant 1, \{a_{i,n}, i \geqslant 0\}$ 是概率分布. 令

$$A(s) = \sum_{i=0}^{\infty} a_i s^i, \quad A_n(s) = \sum_{i=0}^{\infty} a_{i,n} s^i \quad (|s| < 1)$$

是它们的母函数, 试证

$$\lim_{n\to\infty} a_{i,n} = a_i \quad (i \geqslant 0)$$

的充分必要条件是

$$\lim_{n\to\infty} A_n(s) = A(s) \quad (\forall |s| < 1).$$

9. 设 $a_n$ 是 "在一个 $n+1$ 边的凸多边形中, 用 $n-2$ 条不相交的对角线划分为三角形" 的所有可能的方式的数目. 令 $a_1 = 1$. 试证:

$$a_n = a_1 a_{n-1} + a_2 a_{n-2} + \cdots + a_{n-1} a_1 \quad (\forall n \geqslant 2).$$

找出 $\{a_n, n \geqslant 1\}$ 的母函数, 并求出 $a_n$ 的精确表达式.

10. (抽样验收) 假设一批产品中的合格率为 $p \in (0,1)$, 而一件产品被抽出来检验的概率为 $p' \in (0,1)$. 因此, 我们可以把全部产品分成四类: 接收了而且进行了检验的; 接收了而未进行检验的; 抽出来检验了而未接收的; 未抽出来检验且未接收的. 它们的概率分别为 $pp', pq', p'q, qq'$, 其中 $q = 1-p, q' = 1-p'$. 令 $N$ 是发现第一件废品以前已通过检验的产品的件数 (可以是真正检验过的, 也可以是未检验而混过去的), $K$ 为其中没有被发现 (没有检验) 而混过去的废品的件数. 求出 $N$ 和 $K$ 的联合分布及边缘分布. 再求 $\boldsymbol{E}\left(\dfrac{K}{N+1}\right)$ 和 $\operatorname{cov}(K, N)$. (在实际工业生产中, 被发现的废品是要用合格产品去替代的. 因此,$K/(N+1)$ 是其中

废品的比例, 它衡量了这批产品的质量. 注意: $\boldsymbol{E}\left(\dfrac{K}{N+1}\right)$ 与 $\boldsymbol{E}(K)/\boldsymbol{E}(N+1)$ 不一样.)

11. 试证: 若 $X$ 和 $Y$ 这两个随机变量每个都只能取两个数, 而且 $\operatorname{cov}(X,Y)=0$, 则 $X$ 与 $Y$ 相互独立.

12. (样本结构) 设某批产品中有 $r$ 类产品, 它们所占之百分比分别为 $p_1, p_2,\cdots,p_r$ $\left(p_j>0, \sum\limits_{j=1}^{r} p_j=1\right)$. 对其进行有放回的抽样, 假定随机地抽取了一个大小为 $n$ 的样本. 试问样本中不包含那些类的元素的类数的数学期望是多少? (指示: 当第 $j$ 类产生在样本中不出现时令 $X_j=1$, 反之令 $X_j=0$.)

13. (分层抽样) 一个城市有 $n$ 个区, 其中住有 $x_j$ 个居民的区共有 $n_j$ 个 $(n_1+n_2+\cdots=n)$. 令 $m=\sum n_jx_j/n$ 为每个区居民的平均数, 且令 $\sigma^2=\sum(n_jx_j^2/n)-m^2$. 随机地无放回地选取 $r$ 个区, 然后把样本中每一个区的居民数调查出来. 令 $X_1,X_2,\cdots,X_r$ 分别为这 $r$ 个区的居民数. 试证:

$$\boldsymbol{E}(X_1+\cdots+X_r)=rm;$$
$$\operatorname{var}(X_1+\cdots+X_r)=\sigma^2r(n-r)/(n-1).$$

(注意: 有放回的方差较大, 对应的方差为 $r\sigma^2$.)

14. 设 $(\Omega,\mathscr{F},P)$ 是完备概率空间, $(\boldsymbol{B},\|\cdot\|)$ 是可分的 Banach 空间, $X:\Omega\mapsto\boldsymbol{B}$, $X$ 是强可测的 $\boldsymbol{B}$ 值随机元. 定义 4.6 定义了 $X$ 的 (强) 数学期望, 即 Bochner 意义下的数学期望 $\boldsymbol{E}(X)$. 为了区别起见, 这种数学期望有时记之为 (s)$\boldsymbol{E}(X)$. 下面我们定义 Pettis 意义下的数学期望, 或弱数学期望 (w)$\boldsymbol{E}(X)$.

设 $\boldsymbol{B}^*$ 为 $\boldsymbol{B}$ 的共轭空间, 称强可测随机元 $X$ 有 Pettis 意义下的数学期望, 如果存在 $x\in\boldsymbol{B}$, 使得对每一个 $f\in\boldsymbol{B}^*$, $f(X(\omega))$ 关于概率测度 $P$ 是可积的, 而且

$$f(x)=\int_\Omega f(X(\omega))P(d\omega)\quad(\forall f\in\boldsymbol{B}^*),$$

这样的 $x$ 称为 $X$ 在 Pettis 意义下的数学期望, 记之为 (w)$\boldsymbol{E}(X)=x$.

注意: 这样的 $x$ 如果存在, 则一定是唯一的.

试证明: 如果 (s)$\boldsymbol{E}(X)$ 存在, 则 (w)$\boldsymbol{E}(X)$ 也存在且二者相等.

15. (续上) 试列举 (w)$\boldsymbol{E}(X)$ 的一些简单的基本性质并证明之.

16. (续上) 若 (s)$\boldsymbol{E}(X|\mathscr{G})$ 存在 (其中 $\mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, $B_0$ 是 $\boldsymbol{B}$ 的闭子集, 且 $P(X\in\boldsymbol{B}_0)=1$, 则 $P((\mathrm{s})\boldsymbol{E}(X|\mathscr{G})\in\boldsymbol{B}_0)=1$.

17. (续上) 若 $\sigma(X)$ 与 $\mathscr{G}$ 独立, 试证 (s)$\boldsymbol{E}(X|\mathscr{G})=$ (s)$\boldsymbol{E}(X)$.

18. (续上, Jensen 不等式) 设 $(\Omega,\mathscr{F},P)$, $(\boldsymbol{B},\|\cdot\|)$ 如习题 14.

$$p(x):\boldsymbol{B}\mapsto\mathbf{R},$$

称 $p$ 是强凸的, 如果它满足:

(1) $p(x) \geqslant 0, (\forall x \in \boldsymbol{B})$;

(2) $p(\alpha x) = \alpha p(x), (\forall \alpha \geqslant 0, x \in \boldsymbol{B})$;

(3) $p(x+y) \leqslant p(x) + p(y), (\forall x, y \in \boldsymbol{B})$.

试证: (1) 强凸泛函 $p$ 为连续泛函的充分必要条件是: $p$ 是有界的;

(2) (Jensen 不等式) 设 $\mathscr{G}$ 是 $\mathscr{F}$ 中子 $\sigma$ 代数, $\boldsymbol{B}$ 值强可测随机元 $X$ 满足 $\boldsymbol{E}(\|X\|) < \infty$, 则对 $\boldsymbol{B}$ 上的任何连续的强凸的泛函 $p(x)$, 都有:

$$p((\mathrm{s})\boldsymbol{E}(X|\mathscr{G})) \leqslant \boldsymbol{E}(p(X)|\mathscr{G}).$$

# 第三章　各种收敛性

---

## §1　概率收敛、概率为 1 地收敛、$L^p$ 收敛、几乎一致收敛和完全收敛

在这一章中, 将介绍概率论中常用到的一些收敛性. 在 §1 中, 介绍五种收敛性和它们之间的关系.

**定义 1.1**　设 $(\Omega, \mathscr{F}, P)$ 是完备概率空间, $X, X_1, X_2, \cdots$ 是其上的一列实值随机变量.

(1) 若对任意 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} P(|X_n - X| \geqslant \varepsilon) = 0, \tag{1.1}$$

则称 $\{X_n\}$ **概率为 1 地收敛**到 $X$, 记作

$$\lim_{n\to\infty} X_n = X, \quad P, \text{ 或 } X_n \xrightarrow{P} X.$$

(2) 若存在 $\mathscr{F}$ 中的一个 $P-$ 零测集 $A$, 使

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad (\forall \omega \in \Omega - A), \tag{1.2}$$

则称 $\{X_n\}$ 概率为 1 地收敛到 $X$, 记作

$$\lim_{n\to\infty} X_n = X, \quad \text{a.s.}, \quad \text{或} \quad X_n \xrightarrow{\text{a.s.}} X.$$

概率为 1 地收敛即一般测度论中的几乎处处收敛的概率说法.

(3) 如果 $\boldsymbol{E}(|X|^p) < \infty, \boldsymbol{E}(|X_n|^p) < \infty (n = 1, 2, \cdots), p \geqslant 1$, 而且

$$\lim_{n\to\infty} \boldsymbol{E}(|X - X_n|^p) = 0, \tag{1.3}$$

则称 $\{X_n\}L^p-$ **收敛**到 $X$, 记作

$$\lim_{n\to\infty} X_n = X, \quad L^p, \quad 或 \quad X_n \xrightarrow{L^p} X.$$

$L^1-$ 收敛称为平均收敛; $L^2-$ 收敛称为均方收敛.

(4) 如果对任何 $\varepsilon > 0$, 存在 $F_\varepsilon \in \mathscr{F}$, 使 $P(\Omega - F_\varepsilon) < \varepsilon$, 而且

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \ \text{在 } F_\varepsilon \text{ 上一致成立}, \tag{1.4}$$

则称 $\{X_n\}$ **几乎一致收敛**到 $X$, 记作

$$\lim_{n\to\infty} X_n = X, \quad \text{a.un.}, \quad 或 \quad X_n \xrightarrow{\text{a.un.}} X.$$

(5) 如果对任何 $\varepsilon > 0$, 都有

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} P(|X - X_k| \geqslant \varepsilon) = 0, \tag{1.5}$$

则称 $\{X_n\}$ **完全收敛**到 $X$, 记作

$$\lim_{n\to\infty} X_n = X, \quad \text{com.}, \quad 或 \quad X_n \xrightarrow{\text{com.}} X.$$

下面研究这五种收敛性的相互关系.

**定理 1.1** 设 $(\Omega, \mathscr{F}, P), \{X, X_1, X_2, \cdots\}$ 如定义 1.1 中所定义, 则

$$X_n \xrightarrow{\text{com}} X \Rightarrow X_n \xrightarrow{\text{a.s.}} X \Rightarrow X_n \xrightarrow{P} X.$$

**证** 注意: $X_n \xrightarrow{\text{a.s.}} X \Leftrightarrow \forall \varepsilon > 0$ 都有

$$\lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} \{|X - X_k| \geqslant \varepsilon\}\right) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X - X_k| \geqslant \varepsilon\}\right) = 0, \tag{1.6}$$

而且

$$P\left(\bigcup_{k=n}^{\infty} \{|X - X_k| \geqslant \varepsilon\}\right) \leqslant \sum_{k=n}^{\infty} P(|X - X_k| \geqslant \varepsilon), \tag{1.7}$$

所以由 (1.6)、(1.7) 知:

$$X_n \xrightarrow{\text{com.}} X \Rightarrow X_n \xrightarrow{\text{a.s.}} X.$$

由 (1.6) 及概率收敛之定义即得:

$$X_n \xrightarrow{\text{a.s.}} X \Rightarrow X_n \xrightarrow{P} X.$$

定理证毕.

**定理 1.2** 设 $(\Omega, \mathscr{F}, P), \{X, X_1, X_2, \cdots\}$ 如定理 1.1, 则

$$X_n \xrightarrow{\text{a.un.}} X \Rightarrow X_n \xrightarrow{\text{a.s.}} X.$$

**证** 设 $X_n \xrightarrow{\text{a.un.}} X$. 则对每一个 $\varepsilon_k = \dfrac{1}{k}, (k = 1, 2, \cdots)$, 都有 $B_k \in \mathscr{F}$, 使 $P(B_k) \geqslant 1 - \dfrac{1}{k}$ 且 $\lim\limits_{n\to\infty} X_n = X$ 在每个 $B_k$ 上一致成立. 所以对任何 $\omega \in \bigcup\limits_{k=1}^{\infty} B_k$, 都有 $\lim\limits_{n\to\infty} X_n(\omega) = X(\omega)$, 而 $P\left(\bigcup\limits_{k=1}^{\infty} B_k\right) = 1$. 定理证毕.

**定理 1.3** 设 $(\Omega, \mathscr{F}, P), \{X, X_1, X_2, \cdots\}$ 如定理 1.1, 则

$$X_n \xrightarrow{L^p} X \Rightarrow X_n \xrightarrow{P} X.$$

**证** 设 $X_n \xrightarrow{P} X$ 不成立, 则存在一个 $\varepsilon > 0$ 和一列正整数 $\{n_k\}$, 使

$$P(|X - X_{n_k}| \geqslant \varepsilon) \geqslant \delta > 0 \quad (\forall k \geqslant 1).$$

所以

$$\boldsymbol{E}(|X - X_{n_k}|^p) \geqslant \delta \varepsilon^p \quad (\forall k \geqslant 1),$$

从而 $\{X_n\}$ 不可能 $L^p-$ 收敛到 $X$.

**定理 1.4** 设 $(\Omega, \mathscr{F}, P), \{X, X_1, X_2, \cdots\}$ 如定理 1.1, 则

$$\lim_{n\to\infty} X_n = X, \text{a.s.} \Leftrightarrow \lim_{n\to\infty} X_n = X, \text{a.un.}$$

**证** "$\Leftarrow$" 部分在定理 1.2 中已证.

"$\Rightarrow$". 设 $\lim\limits_{n\to\infty} X_n = X$, a.s.. 令

$$B_{k,m} = \left\{|X - X_k| < \frac{1}{m}\right\}, \quad C_{n,m} = \bigcap_{k=n}^{\infty} B_{k,m},$$

则

$$\begin{aligned}
&\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} \\
&= \bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} C_{n,m}, \\
&P\left(\bigcap_{m=1}^{\infty} \bigcup_{n=1}^{\infty} C_{n,m}\right) = 1.
\end{aligned}$$

所以, 对任何 $\varepsilon > 0$ 和任何正整数 $m$, 都存在正整数 $n(\varepsilon, m)$ 使

$$P(C_{n(\varepsilon,m),m}) > 1 - \frac{\varepsilon}{2^m}.$$

取

$$F_\varepsilon = \bigcap_{m=1}^{\infty} C_{n(\varepsilon,m),m} = \bigcap_{m=1}^{\infty} \bigcap_{k=n(\varepsilon,m)}^{\infty} \left\{|X - X_k| < \frac{1}{m}\right\},$$

则 $P(F_\varepsilon) > 1 - \varepsilon$, 且当 $\varepsilon \in F_\varepsilon$ 时, 对任何 $m \geqslant 1$, 都存在 $n(\varepsilon, m)$, 使 $|X(\omega) - X_k(\omega)| < \dfrac{1}{m}$ $(\forall k \geqslant n(\varepsilon, m))$. 这就证明了 $X_n \xrightarrow{\text{a.un.}} X$. 定理证毕.

**定理 1.5** 设 $(\Omega, \mathscr{F}, P), \{X, X_1, X_2, \cdots\}$ 如定理 1.1. 若 $X_n \xrightarrow{P} X$, 则存在正整数的一个子序列 $\{n_k\}$, 使

$$X_{n_k} \xrightarrow{\text{a.s.}} X.$$

**证** 令 $B_{m,n}(\varepsilon) = \{\omega \in \Omega : |X_m(\omega) - X_n(\omega)| \geqslant \varepsilon\}$, 设 $X_n \xrightarrow{P} X$, 则

$$\lim_{\substack{m\to\infty \\ n\to\infty}} P\left(|X_n - X_m| \geqslant \frac{1}{2^k}\right) = 0 \quad (\forall k \geqslant 1).$$

所以, 存在单调上升的正整数列 $\{n_k\}$, 使 $m, n \geqslant n_k$ 时有

$$P\left(B_{m,n}\left(\frac{1}{2^k}\right)\right) < \frac{1}{2^k}, \quad k = 1, 2, \cdots.$$

令

$$B_k = B_{n_k, n_{k+1}}\left(\frac{1}{2^k}\right),$$

则 $P(B_k) < \dfrac{1}{2^k}$. 所以

$$\omega \in \Omega - \bigcup_{m=i}^{\infty} B_m \Rightarrow \left|X_{n_m}(\omega) - X_{n_{m+1}}(\omega)\right| < \frac{1}{2^m} \quad (\forall m \geqslant i).$$

因此, 当 $j \geqslant i$ 且 $\omega \in \bigcup\limits_{m=i}^{\infty} B_m$ 时有

$$\begin{aligned}\left|X_{n_i}(\omega) - X_{n_j}(\omega)\right| &\leqslant \sum_{m=i}^{\infty} \left|X_{n_m}(\omega) - X_{n_{m+1}}(\omega)\right| \\ &\leqslant \sum_{m=i}^{\infty} \frac{1}{2^m} = \frac{1}{2^{i-1}}.\end{aligned}$$

但是

$$P\left(\Omega - \bigcup_{m=i}^{\infty} B_m\right) \geqslant 1 - \frac{1}{2^{i-1}},$$

此即

$$\lim_{\substack{i\to\infty \\ j\to\infty}} \left|X_{n_i} - X_{n_j}\right| = 0, \quad \text{a.un.}.$$

更有

$$\lim_{\substack{i\to\infty \\ j\to\infty}} \left|X_{n_i} - X_{n_j}\right| = 0, \quad \text{a.s.}.$$

由实数空间的完备性, 得知存在实值随机变量 $Y$, 使

$$\lim_{k\to\infty} X_{n_k} = Y, \quad \text{a.s.},$$

但是

$$P(X \neq Y) \leqslant \sum_{m=1}^{\infty} P\left(|X-Y| \geqslant \frac{1}{m}\right),$$

$$P\left(|X-Y| \geqslant \frac{1}{m}\right) \leqslant P\left(|X-X_{n_k}| \geqslant \frac{1}{2m}\right) + P\left(|Y-X_{n_k}| \geqslant \frac{1}{2m}\right).$$

令 $k \to \infty$, 可得 $P\left(|X-Y| \geqslant \dfrac{1}{m}\right) = 0 \ (\forall m \geqslant 1)$, 从而

$$P(X \neq Y) = 0.$$

所以

$$X_{n_k} \xrightarrow{\text{a.s.}} X.$$

定理 1.5 证毕.

**定理 1.6** 定义 1.1 中所定义的 5 种收敛性的极限在 a.s. 的意义下都是唯一的.

**证** 只证第 5 种收敛性的极限在 a.s. 的意义下是唯一的. 其余 4 种作为习题, 请读者自行验证之. 设实值随机变量 $X$ 和 $Y$ 满足

$$P(X \neq Y) > 0,$$

则存在 $\varepsilon > 0$ 和正整数 $M$, 使

$$P\left(|X-Y| \geqslant \frac{1}{m}\right) \geqslant \varepsilon > 0 \quad (\forall m \geqslant M). \tag{1.8}$$

但是,

$$\begin{aligned}
&\sum_{k=1}^{\infty} P\left(|X-X_k| \geqslant \frac{1}{2m}\right)+\sum_{k=1}^{\infty} P\left(|X_k-Y| \geqslant \frac{1}{2m}\right)\\
\geqslant &\sum_{k=1}^{\infty} P\left(\left\{|X-X_k| \geqslant \frac{1}{2m}\right\} \bigcup \left\{|X_k-Y| \geqslant \frac{1}{2m}\right\}\right)\\
\geqslant &\sum_{k=1}^{\infty} P\left(\left\{|X-X_k|+|X_k-Y|>\frac{1}{m}\right\}\right)\\
\geqslant &\sum_{k=1}^{\infty} P\left(|X-Y|>\frac{1}{m}\right)\\
\overset{(1.8)}{=\!=} &\ \infty \qquad (\forall m \geqslant M).
\end{aligned}$$

所以 $\{X_n\}$ 绝不可能完全收敛到两个不 a.s. 相等的随机变量.

**附注 1.1** 完全收敛性蕴涵了 a.s. 收敛性, 但反之不真. 反例如下: 取

$\Omega=[0,1)$,

$\mathscr{F}=[0,1)$ 中的一切 Borel 子集,

$P$ 是 $\mathscr{F}$ 上的 Lebesgue 测度.

定义随机变量列如下

$$X_n(\omega)=\begin{cases}1, & \text{当 } \omega \in \left[0, \dfrac{1}{n}\right), \\ 0, & \text{反之}\end{cases} \qquad (n \geqslant 1).$$

显然 $\lim\limits_{n \to \infty} X_n=0$, a.s.. 但是 $P(|X_n| \geqslant 1)=\dfrac{1}{n}(n \geqslant 1)$, 所以 $\{X_n\}$ 不能完全收敛到 0.

**定义 1.2** 称随机变量 $X$ 与 $Y$ 依分布等价, 如果 $X$ 与 $Y$ 的分布相同. 称两个随机变量序列 $\{X_n, n \geqslant 1\}$ 与 $\{Y_n, n \geqslant 1\}$ **依分布等价**, 如果对每个 $n \geqslant 1$, $X_n$ 与 $Y_n$ 依分布等价.

**定理 1.7** 随机变量列 $\{X_n, n \geqslant 1\}$ 完全收敛到 $X_0$ 的充分必要条件是: 对每一个与 $\{X_n-X_0, n \geqslant 1\}$ 依分布等价的随机变量列 $\{Y_n, n \geqslant 1\}$ 都有

$$\lim_{n \to \infty} Y_n=0, \quad \text{a.s.}.$$

**证** 必要性显然成立. 下面证明充分性. 任取一个相互独立的与 $\{X_n-X_0, n \geqslant 1\}$ 依分布等价的随机变量列 $\{Y_n, n \geqslant 1\}$, 如果它 a.s. 收敛到 0, 则依 a.s. 收敛之定义有:

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty}\{|Y_k| \geqslant \varepsilon\}\right)=0 \quad (\forall \varepsilon>0).$$

但是 $\{|Y_k| \geqslant \varepsilon, k=1,2,\cdots\}$ 是相互独立的, 由 Borel-Contelli 引理知

$$\sum_{n=1}^{\infty} P(|Y_n| \geqslant \varepsilon) < \infty \quad (\forall \varepsilon > 0),$$

由于 $\{Y_n, n \geqslant 1\}$ 与 $\{X_n - X_0, n \geqslant 1\}$ 依分布等价, 所以

$$\sum_{n=1}^{\infty} P(|X_n - X_0| \geqslant \varepsilon) < \infty,$$

此即

$$\lim_{n\to\infty} X_n = X_0, \quad \text{com.}$$

定理证毕.

**推论 1.1** 若 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, 则 $\{X_n\}$ 完全收敛的充分必要条件是: $\{X_n\}$a.s. 收敛.

**附注 1.2** 本节讨论的五种收敛性, 都是对实值随机变量而言. 其实对取值于可分的 Banach 空间 $(\boldsymbol{B}, \|\bullet\|)$ 的随机元而言, 亦可类似地定义这五种收敛性, 并有类似的结果, 只不过把绝对值 ($\mathbf{R}$ 中的范数) 代之以 $\boldsymbol{B}$ 中的范数 $\|\bullet\|$, 把 Lebesgue 积分代之以 Bochner 积分而已. 读者可作为习题给出这五种收敛性的定义, 并讨论它们之间的关系.

## §2 几个不等式

在这一节中, 我们将把概率论中常用的不等式综合论述之.

**引理 2.1** 设 $a, b$ 为非负实数, $\alpha, \beta$ 为正实数且 $\alpha+\beta=1$, 则

$$a^{\alpha} b^{\beta} \leqslant \alpha a + \beta b. \tag{2.1}$$

**证** 当 $ab=0$ 时, (2.1) 式显然成立. 而当 $a>0, b>0$ 时, 令

$$\varphi(t) = \alpha t^{\frac{1}{\alpha}} + \beta t^{-\frac{1}{\beta}} \quad (t>0),$$

对 $\varphi(t)$ 求导数得

$$\varphi'(t) = t^{\frac{1}{\alpha}-1} - t^{-\frac{1}{\beta}-1} \quad (t>0).$$

解 $\varphi'(t)=0 (t>0)$, 得 $t=1$. 由 $\dfrac{1}{\alpha}>1, \dfrac{1}{\beta}>1$ 可知 $\varphi''(t)>0 (\forall t>0)$. 所以 $\varphi(t)$ 在 $t=1$ 达最小值, 即是

$$\alpha t^{\frac{1}{\alpha}} + \beta t^{-\frac{1}{\beta}} = \varphi(t) \geqslant \varphi(1) = \alpha+\beta = 1 (\forall t>0).$$

在上式中令 $t=\left(\frac{a}{b}\right)^{\alpha\beta}$ 得

$$1\leqslant\alpha\left(\frac{a}{b}\right)^{\beta}+\beta\left(\frac{b}{a}\right)^{\alpha}.$$

乘 $a^{\alpha}b^{\beta}$ 于上式两边并注意 $\alpha+\beta=1$ 即得 (2.1). 引理证毕.

在本节中, 恒设 $(\Omega,\mathscr{F},P)$ 为完备概率空间, $X,Y,X_0,X_1,X_2,\cdots$ 为 $\Omega$ 上的实值随机变量.

**定理 2.1 (Chebyshev 不等式)** 设 $\boldsymbol{E}(X^2)<\infty$, 则对任何常数 $a>0$, 恒有

$$P\big(|X-\boldsymbol{E}(X)|\geqslant a\sqrt{\mathrm{var}(X)}\big)\leqslant\frac{1}{a^2}. \tag{2.2}$$

**证** 用方差之定义立即可证此不等式.

**定理 2.2** 设 $g:\mathbf{R}\mapsto\mathbf{R}_+$, 且 $g\in\mathscr{B}(\mathbf{R})/\mathscr{B}(\mathbf{R}_+)$. 约定 $\frac{\alpha}{0}=\infty$, 其中 $\alpha$ 是正实数, $\frac{0}{0}=1$.

(1) 如果 $g$ 是偶函数而且在 $\mathbf{R}_+$ 上单调非降, 则对任何非负实数 $a\geqslant 0$, 都有

$$\frac{\boldsymbol{E}(g(X))-g(a)}{\sup g(X)}\leqslant P(|X|\geqslant a)\leqslant\frac{\boldsymbol{E}(g(X))}{g(a)}; \tag{2.3}$$

(2) 如果 $g$ 在 $\mathbf{R}$ 上单调非降, 则 (2.3) 中间那一项替换为 $P(X\geqslant a)$ 后仍然成立, 这时 $a\in\mathbf{R}$.

**证** (1) 令 $A=\{|X|\geqslant a\}$, 则由

$$\boldsymbol{E}(g(X))=\int_A g(X)dP+\int_{\Omega-A}g(X)dP \tag{2.4}$$

和

$$g(a)P(A)\leqslant\int_A g(X)dP\leqslant(\sup g(X))P(A), \tag{2.5}$$

$$0\leqslant\int_{\Omega-A}g(X)dP\leqslant g(a), \tag{2.6}$$

可得

$$g(a)P(A)\leqslant\boldsymbol{E}(g(X))\leqslant(\sup g(X))P(A)+g(a). \tag{2.7}$$

(2) 仿 (1) 可证 (2) 也成立.

**特例 1** 当 $g(x)=e^{rx}$, 其中 $r>0$ 是常数, 用结论 (2), 对一任何 $a\in\mathbf{R}$, 有

$$\frac{\boldsymbol{E}(e^{rX})-e^{ra}}{\sup e^{rX}}\leqslant P(X\geqslant a)\leqslant e^{-ra}\boldsymbol{E}(e^{rX}). \tag{2.8}$$

**特例 2** 取 $g(x) = |x|^r$, 其中 $r > 0$ 是常数, 用结论 (1), 对任何 $a > 0$ 有

$$\frac{\boldsymbol{E}(|X|^r) - a^r}{\sup |X|^r} \leqslant P(|X| \geqslant a) \leqslant \frac{\boldsymbol{E}(|X|^r)}{a^r}. \tag{2.9}$$

(2.9) 式右边的不等式称为 Markov 不等式, 而当 $r = 2, \boldsymbol{E}(X) = 0$ 时, (2.9) 式右边的不等式就是 Chebyshev 不等式.

**定理 2.3 (Hölder 不等式)** 设 $p > 1, \dfrac{1}{p} + \dfrac{1}{q} = 1, \boldsymbol{E}(|X|^p) < \infty, \boldsymbol{E}(|Y|^q) < \infty$, 则

$$\boldsymbol{E}(|XY|) < \infty, \tag{2.10}$$

而且对任何 $B \in \mathscr{F}$ 有:

$$\int_B |XY| dP \leqslant \left(\int_B |X|^p dP\right)^{\frac{1}{p}} \left(\int_B |Y|^q dP\right)^{\frac{1}{q}}. \tag{2.11}$$

**证** 若 $|X(\omega)| \leqslant |Y(\omega)|^{q-1}$, 则 $|X(\omega)Y(\omega)| \leqslant |Y(\omega)|^q$; 若 $|X(\omega)| > |Y(\omega)|^{q-1}$, 则 $|X(\omega)Y(\omega)| \leqslant |X(\omega)|^{1+\frac{1}{q-1}} = |X(\omega)|^p$. 总之, 由 $\boldsymbol{E}(|X|^p) < \infty$ 和 $\boldsymbol{E}(|Y|^q) < \infty$ 得知 $\boldsymbol{E}(|XY|) < \infty$.

现在证明不等式 (2.11). 若 $X$ 和 $Y$ 至少有一个在 $B$ 上 a.s. 为 0, 则 (2.11) 显然成立. 所以不失普遍性可设

$$\int_B |X|^p dP > 0, \quad \int_B |Y|^q dP > 0.$$

取

$$\alpha = \frac{1}{p}, \quad \beta = \frac{1}{q}, \quad a = \frac{|X(\omega)|^p}{\displaystyle\int_B |X|^p dP}, \quad b = \frac{|Y(\omega)|^q}{\displaystyle\int_B |Y|^q dP},$$

则由 (2.1) 知: 对每个 $\omega \in B$, 总有

$$\begin{aligned} &\frac{|X(\omega)Y(\omega)|}{\left(\displaystyle\int_B |X|^p dP\right)^{\frac{1}{p}} \left(\displaystyle\int_B |Y|^q dP\right)^{\frac{1}{q}}} \\ \leqslant\ & \frac{|X(\omega)|^p}{p\displaystyle\int_B |X|^p dP} + \frac{|Y(\omega)|^q}{q\displaystyle\int_B |Y|^q dP}. \end{aligned} \tag{2.12}$$

把 (2.12) 两边在 $B$ 上积分并注意 $\dfrac{1}{p} + \dfrac{1}{q} = 1$ 即得 (2.11). 定理证毕.

**附注 2.1** 当 $p = q = 2$ 时, (2.11) 称为 Schwartz 不等式.

**推论 2.1** 当 $1 \leqslant p_1 < p_2$ 时,

$$X_n \xrightarrow{L^{p_2}} X_0 \Rightarrow X_n \xrightarrow{L^{p_1}} X_0. \tag{2.13}$$

**证** 在 (2.11) 中令 $X = |X_n - X_0|^{p_1}, p = \dfrac{p_2}{p_1}, Y \equiv 1, B = \Omega$, 则有

$$\int_\Omega |X_n - X_0|^{p_1} dP \leqslant \left(\int_\Omega |X_n - X_0|^{p_2} dP\right)^{\frac{p_1}{p_2}},$$

所以 (2.13) 成立.

**推论 2.2** 若对某个 $p > 1$, 有

$$\sum_{n=1}^\infty X_n = X_0, \quad L^p,$$

而且 $\boldsymbol{E}(|Y|^q) < \infty, \dfrac{1}{p} + \dfrac{1}{q} = 1$, 则

$$\sum_{n=1}^\infty X_n Y = X_0 Y, \quad L^1. \tag{2.14}$$

**证** 由 (2.11), 对每个 $k \geqslant 1$ 都有

$$\begin{aligned}&\int_\Omega \left|\sum_{n=1}^k X_n Y - X_0 Y\right| dP\\ \leqslant &\left(\int_\Omega \left|\sum_{n=1}^k X_n - X\right|^p dP\right)^{\frac{1}{p}} \left(\int_\Omega |Y|^q dP\right)^{\frac{1}{q}}.\end{aligned}$$

在上式中令 $k \to \infty$ 即得 (2.14).

**定理 2.4 (Minkowski 不等式)** 设对某个 $p \geqslant 1, \boldsymbol{E}(|X|^p) < \infty, \boldsymbol{E}(|Y|^p) < \infty$, 则 $\boldsymbol{E}(|X+Y|^p) < \infty$, 而且对任何 $B \in \mathscr{F}$, 有

$$\left(\int_B |X+Y|^p dP\right)^{\frac{1}{p}} \leqslant \left(\int_B |X|^p dP\right)^{\frac{1}{p}} + \left(\int_B |Y|^p dP\right)^{\frac{1}{p}}. \tag{2.15}$$

**证** $\boldsymbol{E}(|X+Y|^p) < \infty$ 的证明仿定理 2.3. 下面证明 (2.15). 当 $p = 1$, (2.15) 显然成立. 当 $p > 1$ 时, 有

$$\int_B |X+Y|^p dP \leqslant \int_B |X||X+Y|^{p-1} dP + \int_B |Y||X+Y|^{p-1} dP. \tag{2.16}$$

应用 (2.11) 于 (2.16) 右端的两个积分 (取 $q=\frac{p}{p-1}$, 于是 $\frac{1}{p}+\frac{1}{q}=1$) 得

$$\int_B |X+Y|^p dP \leqslant \left(\int_B |X|^p dP\right)^{\frac{1}{p}} \left(\int_B |X+Y|^p dP\right)^{\frac{1}{q}} + \left(\int_B |Y|^p dP\right)^{\frac{1}{p}} \left(\int_B |X+Y|^p dP\right)^{\frac{1}{q}}. \tag{2.17}$$

若 $\int_B |X+Y|^p dP=0$, 则 (2.15) 显然成立. 反之, 用 $\left(\int_B |X+Y|^p dP\right)^{\frac{1}{q}}$ 除 (2.17) 的两边, 并注意 $1-\frac{1}{q}=\frac{1}{p}$ 即得 (2.15). 定理证毕.

**定理 2.5 (Lévy 不等式)** 设 $\{X_k\}$ 为独立随机变量列, $S_n=\sum_{k=1}^n X_k, \mu(S_n)$ 表 $S_n$ 的中位数, 则对任意 $\varepsilon>0$ 有:

(1) $P\left(\max_{k\leqslant n}(S_k-\mu(S_k-S_n))\geqslant\varepsilon\right)\leqslant 2P(S_n\geqslant\varepsilon)$;

(2) $P\left(\max_{k\leqslant n}|S_k-\mu(S_k-S_n)|\geqslant\varepsilon\right)\leqslant 2P(|S_n|\geqslant\varepsilon)$.

证明甚易, 读者可作为习题验证之或参见 [90] p.260.

**推论 2.3** 设 $\{X_k\}$ 是数学期望为 0 方差有限的独立随机变量列, $S_n=\sum_{k=1}^n X_k$, 则对任意 $\varepsilon>0$, 有

$$P\left(\max_{k\leqslant n} S_k\geqslant\varepsilon\right)\leqslant 2P(S_n\geqslant\varepsilon-\sqrt{2\mathrm{var}(S_n)}).$$

## §3 弱 收 敛

在 §1 中, 我们讨论了五种收敛性, 它们都是关于 (点) 函数列的收敛性. 在第 3 节, 我们将要研究测度 (集函数) 列的收敛性.

在这一节中, 我们采用第一章第 3 节的符号. 设 $(\boldsymbol{B},\rho)$ 是一个距离空间, $\mathscr{G}$ 是由距离 $\rho$ 所产生的拓扑 (全体开集所成之集合系), $\mathscr{B}=\mathscr{B}(\boldsymbol{B})=\sigma(\mathscr{G})$ 是由 $\mathscr{G}$ 所产生的 $\sigma$ 代数, 称之为 Borel $\sigma$ 代数, $\mathscr{B}(\boldsymbol{B})$ 中每个集合 $B$ 都称为 Borel 集. 于是得到一个 Borel 可测距离空间 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$. 再令 $C_b(\boldsymbol{B})$ 为定义在 $\boldsymbol{B}$ 上的全体有界连续实值函数.$C_{b,u}(\boldsymbol{B})$ 为 $C_b(\boldsymbol{B})$ 中全体一致连续函数.

**定义 3.1** 设 $\{P,P_1,P_2,\cdots\}$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的一列概率测度, 如果对任意 $f\in C_b(\boldsymbol{B})$, 都有

$$\lim_{n\to\infty}\int_B f dP_n=\int_B f dP, \tag{3.1}$$

则称 $\{P_n, n \geqslant 1\}$ **弱收敛**到 $P$, 记之为

$$\lim_{n\to\infty} P_n = P, \quad \text{w.,} \quad 或 \quad P_n \xrightarrow{\text{w}} P.$$

**定理 3.1** 设 $F$ 是 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 中任一闭集. 对任何 $\varepsilon > 0$, 都存在 $f_\varepsilon \in C_b(\boldsymbol{B})$, 使 $0 \leqslant f_\varepsilon(x) \leqslant 1, (\forall x \in \boldsymbol{B})$, 而且

$$f_\varepsilon(x) = \begin{cases} 1, & 当 x \in F, \\ 0, & 当 \rho(x, F) \geqslant \varepsilon, \end{cases} \tag{3.2}$$

其中 $\rho(x, F) = \inf\{\rho(x, y) : y \in F\}$ 为 $x$ 到 $F$ 的距离. $f_\varepsilon$ 可取成一致连续的.

**证** 令

$$\varphi(t) = \begin{cases} 1, & 当 t \leqslant 0, \\ 1 - t, & 当 0 \leqslant t \leqslant 1, \\ 0, & 当 1 \leqslant t, \end{cases} \tag{3.3}$$

取

$$f_\varepsilon(x) = \varphi\left(\frac{1}{\varepsilon}\rho(x, F)\right), \tag{3.4}$$

则 $f_\varepsilon$ 即为所求.

定理 3.1 说明 $\mathbf{1}_F$ 可以用一致连续的有界函数列来逼近.

**定理 3.2** 设 $P$ 和 $Q$ 都是 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 上的概率测度, 如果对任何 $f \in C_b(\boldsymbol{B})$, 都有

$$\int_{\boldsymbol{B}} f dP = \int_{\boldsymbol{B}} f dQ, \tag{3.5}$$

则 $P$ 和 $Q$ 在 $\mathscr{B}(\boldsymbol{B})$ 上恒等.

在证明定理 3.2 以前, 先证明

**引理 3.1** $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 上的任何一个概率测度 $P$ 都是正则的, 即是对任何 Borel 集 $A \in \mathscr{B}(\boldsymbol{B})$, 任何 $\varepsilon > 0$, 都存在闭集 $F$ 与开集 $G$, 使 $F \subset A \subset G$, 且 $P(G - F) < \varepsilon$.

**证** 令 $\mathfrak{M}$ 为具有引理 3.1 中的性质的 Borel 集合的全体. 设 $A$ 是闭集, 取 $F = A, G_\delta = \{x \in \boldsymbol{B} : \rho(x, A) < \delta\}$, 则开集 $G_\delta$ 当 $\delta \downarrow 0$ 时单调下降到 $A$. 这说明任何闭集均属于 $\mathfrak{M}$. 若还能证明 $\mathfrak{M}$ 是一个 $\sigma$ 代数, 则引理 3.1 获证. 事实上, 任取 $\{A_n\} \subset \mathfrak{M}$, 取闭集列 $\{F_n\}$ 和开集列 $\{G_n\}$ 使

$$F_n \subset A_n \subset G_n, \quad P(G_n - F_n) \leqslant \varepsilon/2^{n+1} \quad (\forall n \geqslant 1).$$

再取开集 $G = \bigcup\limits_{n\geqslant 1} G_n$ 和闭集 $F = \bigcup\limits_{n\leqslant n_0} F_n$, 使

$$P\left(\bigcup_{n\geqslant 1} F_n - F\right) \leqslant \varepsilon/2, \quad 则 \quad F \subset \bigcup_{n\geqslant 1} A_n \subset \bigcup_{n\geqslant 1} G_n,$$

且 $P(G-F) \leqslant \varepsilon$. 这说明 $\mathfrak{M}$ 中的集合对可数并运算是封闭的. 显然 $\mathfrak{M}$ 中的集合对补的运算也是封闭的. 所以 $\mathfrak{M}$ 是一个 $\sigma$ 代数. 引理证毕.

**附注 3.1** 对全空间上测度值有限之测度, 引理亦成立.

下面用引理 3.1 来证明定理 3.2.

设 $F$ 是 $\boldsymbol{B}$ 中的闭集, $\varphi$ 如 (3.3) 中所定义. 再令

$$\varphi_u(t) = \varphi(ut), \tag{3.6}$$

$$f_u(x) = \varphi_u(\rho(x, F)). \tag{3.7}$$

则 $\{f_u\} \subset C_b(\boldsymbol{B}), |f_u| \leqslant 1$, 且 $\lim\limits_{u\to\infty} f_u(x) = \mathbf{1}_F(x)$. 用 Lebesgue 控制收敛定理和 (3.5) 得

$$P(F) = \lim_{u\to\infty} \int_{\boldsymbol{B}} f_u dP \overset{(3.5)}{=\!=} \lim_{u\to\infty} \int_{\boldsymbol{B}} f_u dQ = Q(F).$$

这说明 $P$ 和 $Q$ 在 $\Omega$ 中一切闭集均相等. 再用引理 3.1 知 $P$ 和 $Q$ 在 $\mathscr{B}(\boldsymbol{B})$ 上恒等. 定理证毕.

**推论 3.1** 弱收敛的极限是唯一的.

**定理 3.3 (弱收敛的等价描述)** 设 $\{P, P_1, P_2, \cdots\}$ 是 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 上的一列概率测度, 则下列五种陈述等价:

(1) $P_n \overset{\text{w}}{\Longrightarrow} P$;

(2) $\lim\limits_{n\to\infty} \int_{\boldsymbol{B}} f dP_n = \int_{\boldsymbol{B}} f dP \quad (\forall f \in C_{b,u}(\boldsymbol{B}))$;

(3) $\limsup\limits_{n\to\infty} P_n(F) \leqslant P(F)$ (对一切闭集 $F$);

(4) $\liminf\limits_{n\to\infty} P_n(G) \geqslant P(G)$ (对一切开集 $G$);

(5) $\lim\limits_{n\to\infty} P_n(A) = P(A)$ (对一切 $P$ 连续集 $A$).

所谓 $A$ 是 $P$ 连续集, 意即 $P(A^\circ) = P(A) = P(\overline{A})$, 其中 $A^\circ$ 是 $A$ 之开核, $\overline{A}$ 是 $A$ 之闭包.

**证** 我们证明此定理的逻辑思路是

$$\begin{array}{c}(1) \Rightarrow (2) \Rightarrow (3) \Leftrightarrow (4); (3) \Rightarrow (1).\\ \qquad\qquad\qquad\quad\ \Updownarrow \qquad\qquad\qquad\\ \qquad\qquad\qquad\quad\ (5) \qquad\qquad\qquad\end{array}$$

(1) $\Rightarrow$ (2) 是显然的.

(2) $\Rightarrow$ (3). 设 (2) 成立, 而且 $F$ 是闭集. 任给 $\delta > 0$, 选足够小的 $\varepsilon$ 及 $G_\varepsilon = \{x \in \boldsymbol{B}; \rho(x, F) < \varepsilon\}$, 使 $P(G_\varepsilon) < P(F) + \delta$. 由于 $\{G_\varepsilon, \varepsilon > 0\}$ 当 $\varepsilon \downarrow 0$ 时单调下降到 $F$, 所以若 $f_\varepsilon$ 由 (3.4) 所定义, 则有: $f_\varepsilon$ 在 $\boldsymbol{B}$ 上一致连续, $f_\varepsilon(x) = 1(\forall x \in F), f_\varepsilon(x) = 0(\forall x \in G_\varepsilon^c)$ 且 $0 \leqslant f_\varepsilon \leqslant 1$. 由于 (2) 成立, 所以

$$\lim_{n\to\infty} \int_{\boldsymbol{B}} f_\varepsilon dP_n = \int_{\boldsymbol{B}} f_\varepsilon dP. \tag{3.8}$$

又因为

$$P_n(F) = \int_F f_\varepsilon dP_n \leqslant \int_{\boldsymbol{B}} f_\varepsilon dP_n, \tag{3.9}$$

$$\int_{\boldsymbol{B}} f_\varepsilon dP = \int_{G_\varepsilon} f_\varepsilon dP \leqslant P(G_\varepsilon) < P(F) + \delta. \tag{3.10}$$

由 (3.8)~(3.10) 得

$$\begin{aligned}\limsup_{n\to\infty} P_n(F) &\leqslant \lim_{n\to\infty} \int_{\boldsymbol{B}} f_\varepsilon dP_n \\ &\overset{(2)}{=\!=} \int_{\boldsymbol{B}} f_\varepsilon dP \leqslant P(F) + \delta.\end{aligned} \tag{3.11}$$

由于 $\delta > 0$ 可以任意小, 所以由 (3.11) 可知 (3) 成立.

(3) $\Rightarrow$ (1). 设 (3) 成立. 任取 $f \in C_b(\boldsymbol{B})$. 我们首先证明:

$$\limsup_{n\to\infty} \int_{\boldsymbol{B}} f dP_n \leqslant \int_{\boldsymbol{B}} f dP. \tag{3.12}$$

不失普遍性可设 $0 < f(x) < 1(\forall x \in \boldsymbol{B})$. (否则, 若 $f \equiv 0$, 则 (3.12) 显然成立; 若 $K \overset{\text{def.}}{=\!=} \sup\limits_{x\in\boldsymbol{B}} |f(x)| > 0$, 定义 $f$ 的线性变换 $\bar{f} \overset{\text{def.}}{=\!=} \dfrac{1}{6K} f + \dfrac{2}{3}$ 取代 $f$.)

暂时固定正整数 $k$, 令闭集

$$F_i = \left\{x \in \boldsymbol{B} : f(x) \geqslant \frac{i}{k}\right\}, \quad i = 0, 1, \cdots, k.$$

因为 $0 < f < 1$, 所以

$$\sum_{i=1}^k \frac{i-1}{k} P\left(\frac{i-1}{k} \leqslant f < \frac{i}{k}\right) \leqslant \int_{\boldsymbol{B}} f dP \leqslant \sum_{i=1}^k \frac{i}{k} P\left(\frac{i-1}{k} \leqslant f < \frac{i}{k}\right). \tag{3.13}$$

若注意 $F_0 = \Omega, F_k = \varnothing$, 则可知 (3.13) 右端等于

$$\sum_{i=1}^k \frac{i}{k} [P(F_{i-1}) - P(F_i)] = \frac{1}{k} + \frac{1}{k} \sum_{i=1}^k P(F_i). \tag{3.14}$$

仿之, (3.13) 左端的和等于

$$\frac{1}{k}\sum_{i=1}^{k}P(F_i). \tag{3.15}$$

由 (3.13)—(3.15) 得

$$\frac{1}{k}\sum_{i=1}^{k}P(F_i)\leqslant\int_{\boldsymbol{B}}fdP\leqslant\frac{1}{k}+\sum_{i=1}^{k}P(F_i). \tag{3.16}$$

注意: 在 (3.16) 的推导中并不要求两个不等式中的概率测度都是一样的 $P$. 实际上下述两个不等式亦成立:

$$\frac{1}{k}\sum_{i=1}^{k}P(F_i)\leqslant\int_{\boldsymbol{B}}fdP;\quad\int_{\boldsymbol{B}}fdP_n\leqslant\frac{1}{k}+\frac{1}{k}\sum_{i=1}^{k}P_n(F_i).$$

所以由 $\limsup\limits_{n\to\infty}P_n(F_i)\leqslant P(F_i)(i=0,1,\cdots k)$ 得

$$\begin{aligned}\limsup_{n\to\infty}\int_{\boldsymbol{B}}fdP_n&\leqslant\frac{1}{k}+\frac{1}{k}\sum_{i=1}^{k}P(F_i)\\&\leqslant\frac{1}{k}+\int_{\boldsymbol{B}}fdP.\end{aligned} \tag{3.17}$$

在 (3.17) 中令 $k\to\infty$ 即得 (3.12).

应用 (3.12) 于 $-f$ 得

$$\liminf_{n\to\infty}\int_{\boldsymbol{B}}fdP_n\geqslant\int fdP\quad(\forall f\in C_b(\boldsymbol{B})). \tag{3.18}$$

由 (3.12), (3.18) 得 $P_n\stackrel{\mathrm{w}}{\Longrightarrow}P$.

(3) ⇔ (4). 由开集之补集是闭集立即可得.

(3) ⇒ (5). 设 (3) 成立, 从而 (4) 亦成立. 对任何 $A\in\mathscr{B}(\boldsymbol{B})$, 恒有

$$\begin{aligned}P(\overline{A})&\geqslant\limsup_{n\to\infty}P_n(\overline{A})\geqslant\limsup_{n\to\infty}P_n(A)\\&\geqslant\liminf_{n\to\infty}P_n(A)\geqslant\liminf_{n\to\infty}P_n(A^\circ)\geqslant P(A^\circ).\end{aligned} \tag{3.19}$$

若 $A$ 是 $P$ **连续集**, 即是 $A$ 的边界 $\partial A\stackrel{\text{def.}}{=\!=}\overline{A}-A^\circ$ 的 $P$ 测度 $P(\partial A)=0$, 所以 $P(\overline{A})=P(A^\circ)$, 由 (3.19) 和 (5) 成立.

(5) ⇒ (3). 设 (5) 成立. 对任何闭集 $F$, 边界 $\partial\{x\in\Omega:\rho(x,F)\leqslant\delta\}$ 含于 $\{x\in\boldsymbol{B}:\rho(x,F)=\delta\}$. 这些边界对不同的 $\delta$ 是两两不交的, 所以至多只有可数多个具有 $P$ 正测度, 所以存在一列趋于 0 的正数 $\delta_k\downarrow 0$, 使 $F_k\stackrel{\text{def.}}{=\!=}\{x\in\Omega:\rho(x,F)\leqslant\delta_k\}$ 是 $P$ 连续集 $(\forall k\geqslant 1)$. 由于 (5) 成立, 所以

$$\limsup_{n\to\infty}P_n(F)\leqslant\lim_{n\to\infty}P_n(F_k)=P(F_k)\quad(\forall k\geqslant 1).$$

由 $F$ 是闭集, 所以 $F_k \downarrow F$. 在上式中令 $k \to \infty$ 即得 (3). 定理证毕.

下面我们介绍随机元列 $\{X_n, n \geqslant 1\}$ 的依分布收敛性.

设 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 仍是前面的可测距离空间, $(\Omega, \mathscr{F}, P)$ 是完备概率空间, $\{X, X_n, n \geqslant 1\}$ 是其上的 $\boldsymbol{B}$ 值随机元列.

**定义 3.2** 称 $\boldsymbol{B}$ 值随机元列 $\{X_n, n \geqslant 1\}$ 依分布收敛到 $\boldsymbol{B}$ 值随机元 $X$, 记之为

$$\lim_{n\to\infty} X_n = X, \ \ d., \ \ 或 \ \ X_n \xrightarrow{d} X,$$

如果 $X_n$ 的 $P$ 分布 $P \circ X_n^{-1}$ 弱收敛到 $X$ 的 $P$ 分布, 即

$$P \circ X_n^{-1} \stackrel{\mathrm{w}}{\Longrightarrow} P \circ X^{-1}.$$

注意: $P \circ X_n^{-1}, P \circ X^{-1}$ 都是 $\mathscr{B}(\boldsymbol{B})$ 上的概率测度.

**定义 3.3** 称集合 $A \in \mathscr{B}(\boldsymbol{B})$ 是 $\boldsymbol{B}$ 值随机元 $X$ 的连续集, 如果 $P(X \in \partial A) = 0$.

类似于定理 3.1, 关于 $\boldsymbol{B}$ 值随机元列的依分布收敛有下面的

**定理 3.4** 设 $\{X, X_n, n \geqslant 1\}$ 是 $(\Omega, \mathscr{F}, P)$ 上一列 $\boldsymbol{B}$ 值随机元, 则下列陈述等价:

(1) $X_n \xrightarrow{\mathrm{d}} X$;

(2) $\lim\limits_{n\to\infty} \boldsymbol{E}(f(X_n)) = \boldsymbol{E}(f(X)) \ \ (\forall f \in C_{b,u}(\boldsymbol{B}))$;

(3) $\limsup\limits_{n\to\infty} P(X_n \in F) \leqslant P(X \in F)$ (对一切闭集 $F$);

(4) $\liminf\limits_{n\to\infty} P(X_n \in G) \geqslant P(X \in G)$ (对一切开集 $G$);

(5) $\lim\limits_{n\to\infty} P(X_n \in A) = P(X \in A)$ (对一切 $X$ 连续集 $A$).

**证** 由定理 3.3 及随机元列依分布收敛的定义即可得定理 3.4.

**附注 3.2** 设 $\{X_n, n \geqslant 1\}$ 是一列 $\boldsymbol{B}$ 值随机元, $Q$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的一个概率测度, 当 $X_n$ 的分布 $P \circ X_n^{-1}$ 弱收敛到 $Q$, 即 $P \circ X_n^{-1} \stackrel{\mathrm{w}}{\Longrightarrow} Q$, 我们有时亦说 $\{X_n\}$ 依分布收敛到 $Q$, 记作

$$X_n \xrightarrow{\mathrm{d}} Q.$$

一个重要的特例是 $\boldsymbol{B} = \mathrm{R}^N$, 这时 $\mathscr{B}(\mathbf{R}^N)$ 是 $\mathbf{R}^N$ 中的 Borel 集全体, 而 $\rho(x,y) = \sqrt{\sum\limits_{i=1}^{N}(x_i - y_i)^2}$ 是 $N$ 维欧氏距离, 其中 $x = (x_1, \cdots, x_N), y = (y_1, \cdots, y_N)$. 而 $\boldsymbol{B}$ 值随机元就是坐标为实值的 $N$ 维随机向量 $X = (X_1, \cdots, X_N)$, 它具有联合分布函数 $F(x_1, \cdots, x_N) \stackrel{\text{def.}}{=} P(X_1 \leqslant x_1, \cdots, X_N \leqslant x_N)$, 这是一个标准的 $(\text{L-S})_N$ 函数, 它产生 $(\text{L-S})_N$ 测度 $\mu_F$.

设 $\{X^{(n)}=(X_1^{(n)},\cdots,X_N^{(n)}),n\geqslant 1\}$ 是一列 $\mathbf{R}^N$ 值随机元, 它有相应的一列联合分布函数

$$F^{(n)}(x_1,\cdots,x_N)=P(X_1^{(n)}\leqslant x_1,\cdots,X_N^{(n)}\leqslant x_N)\quad(n\geqslant 1),$$

和相应的一列 $(\text{L-S})_N$ 测度

$$\{\mu_{F^{(n)}},n\geqslant 1\}.$$

我们亦可讨论 $\{\mu_{F^{(n)}},n\geqslant 1\}$ 的弱收敛和 $\{X^{(n)},n\geqslant 1\}$ 的依分布收敛, 自然有定理 3.3 和 3.4 的特例形式的相应的结果.

**附注 3.3** 定义 3.1 中定义的弱收敛性是对概率测度来说的. 事实上, 若令 $M$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的有限测度 $\mu$ (即 $\mu(\boldsymbol{B})<\infty$) 的全体, 在 $M$ 内的测度列 $\{\mu_n,n\geqslant 1\}$ 亦可完全类似地定义弱收敛, 而且亦有类似的定理 3.2, 定理 3.3 和推论 3.1. 例如

**定理 3.3′** 设 $\{\mu,\mu_n,n\geqslant 1\}\subset M$, 则下列陈述等价:

(1) $\mu_n\stackrel{\text{w}}{\Longrightarrow}\mu$ (定义为 $\int_{\boldsymbol{B}}fd\mu_n\to\int_{\boldsymbol{B}}fd\mu,\forall f\in C_b$);

(2) $\int_{\boldsymbol{B}}fd\mu_n\to\int_{\boldsymbol{B}}fd\mu\quad(\forall f\in C_{b,u})$;

(3) $\mu_n(\boldsymbol{B})\to\mu(\boldsymbol{B})$ 且 $\limsup\limits_{n\to\infty}\mu_n(F)\leqslant\mu(F)$ (对一切闭集 $F$);

(4) $\mu_n(\boldsymbol{B})\to\mu(\boldsymbol{B})$ 且 $\liminf\limits_{n\to\infty}\mu_n(G)\geqslant\mu(G)$ (对一切开集 $G$);

(5) $\lim\limits_{n\to\infty}\mu_n(A)=\mu(A)$, (对一切 $\mu$ 连续集 $A$).

**推论 3.1′** 设 $\{\mu,\nu,\mu_n,n\geqslant 1\}\subset M,\mu_n\stackrel{\text{w}}{\Longrightarrow}\mu,\mu_n\stackrel{\text{w}}{\Longrightarrow}\nu$, 则 $\mu$ 与 $\nu$ 在 $\mathscr{B}(\boldsymbol{B})$ 上恒等.

## §4 局部弱收敛与淡收敛

如 §3 一样, 令 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$ 为可测距离空间, $\mathscr{B}(\boldsymbol{B})$ 是 $\boldsymbol{B}$ 中全体 Borel 集, $\rho$ 是 $\boldsymbol{B}$ 上的距离. $C_b(\boldsymbol{B})$ 为定义在 $\boldsymbol{B}$ 上的全体有界连续的实值函数, 简记 $C_b(\boldsymbol{B})$ 为 $C_b$, $C_{b,b}$ 是 $C_b$ 中的具有有界支撑的函数的全体, $C_{b,k}$ 是 $C_b$ 中具有紧支撑的函数的全体. (实值函数 $f$ 的支撑定义为 $\operatorname{supp}f\stackrel{\text{def.}}{=}\overline{\{f\neq 0\}}$.) 简记 $\mathscr{B}(\boldsymbol{B})=\mathscr{B}$, 而令 $\mathscr{B}_b$ 为 $\mathscr{B}$ 中有界集全体, $\mathscr{B}_k$ 为 $\mathscr{B}$ 中相对紧集全体. 于是有

$$\mathscr{B}_k\subset\mathscr{B}_b\subset\mathscr{B},$$

$$C_{b,k}\subset C_{b,b}\subset C_b.$$

再令 $M_k$ 是定义在 $\mathscr{B}$ 上而在 $\mathscr{B}_k$ 上有限的测度的全体; $M_b$ 是定义在 $\mathscr{B}$ 上而在 $\mathscr{B}_b$ 上有限的测度的全体, $M$ 是定义在 $\mathscr{B}$ 上的有限测度的全体, 于是有:

$$M_k\supset M_b\supset M.$$

**定义 4.1** 设 $\{\mu, \mu_n, n \geqslant 1\} \subset M_b$, 若

$$\lim_{n\to\infty} \int_{\boldsymbol{B}} f d\mu_n = \int_{\boldsymbol{B}} f d\mu, \quad (\forall f \in C_{b,b}), \tag{4.1}$$

则称 $\{\mu_n\}$ **局部弱收敛**到 $\mu$, 记作

$$\lim_{n\to\infty} \mu_n = \mu, \quad \text{lw.},$$

或 $\mu_n \overset{\text{lw}}{\Longrightarrow} \mu$.

设 $\{\mu, \mu_n, n \geqslant 1\} \subset M_k$, 若

$$\lim_{n\to\infty} \int_{\boldsymbol{B}} f d\mu_n = \int_{\boldsymbol{B}} f d\mu, \quad f \in C_{b,k}, \tag{4.2}$$

则称 $\{\mu_n\}$ **淡收敛**到 $\mu$, 记作

$$\lim_{n\to\infty} \mu_n = \mu, \quad \text{v.},$$

或 $\mu_n \overset{\text{v}}{\Longrightarrow} \mu$.

由 $M_k \supset M_b \supset M$ 及 $C_{b,k} \subset C_{b,b} \subset C_b$ 可知: 当 $\{\mu, \mu_n, n \geqslant 1\} \subset M$ 时, "$\mu_n \overset{\text{w}}{\Longrightarrow} \mu$"$\Rightarrow$"$\mu_n \overset{\text{lw}}{\Longrightarrow} \mu$"$\Rightarrow$"$\mu_n \overset{\text{v}}{\Longrightarrow} \mu$".

**定理 4.1 (局部弱收敛的等价描述)** 设 $\{\mu, \mu_n, n \geqslant 1\} \subset M_b$, 则下列陈述等价:

(1) $\mu_n \overset{\text{lw}}{\Longrightarrow} \mu$;

(2) 对任何有界闭集 $F$, 有

$$\limsup_{n\to\infty} \mu_n(F) \leqslant \mu(F), \tag{4.3}$$

对任何有界开集 $G$, 有

$$\liminf_{n\to\infty} \mu_n(G) \geqslant \mu(G); \tag{4.4}$$

(3) 对任何有界 $\mu$ 连续集 $A$, 有

$$\lim_{n\to\infty} \mu_n(A) = \mu(A). \tag{4.5}$$

**证** (1) $\Rightarrow$ (2). 设 (1) 成立, $F$ 是任一有界闭集. 取 $\varphi(t)$ 如 (3.3) 式所定义, 再取

$$G_m = \left\{x \in \boldsymbol{B} : \rho(x, F) < \frac{1}{m}\right\} \quad (m \geqslant 1), \tag{4.6}$$

$$f_m(x) = \varphi(m\rho(x, F)) \quad (m \geqslant 1), \tag{4.7}$$

则有界开集列 $G_m \downarrow F, f_m \in C_{b,b}$, 且

$$\mathbf{1}_F \leqslant f_m \leqslant \mathbf{1}_{G_m} \quad (m \geqslant 1). \tag{4.8}$$

所以, 由 $\mu_n \overset{\text{lw}}{\Longrightarrow} \mu$ 和 (4.8) 式知

$$\begin{aligned}\limsup_{n\to\infty} \mu_n(F) &\leqslant \lim_{n\to\infty} \int_{\boldsymbol{B}} f_m d\mu_n \\ &= \int_{\boldsymbol{B}} f_m d\mu \leqslant \mu(G_m). \end{aligned}\tag{4.9}$$

在 (4.9) 中令 $m \to \infty$ 即得 (4.3).

再设 $G$ 是任一有界开集, 取 $F_m = \left\{x \in \boldsymbol{B} : \rho(x, G^{\mathrm{c}}) \geqslant \dfrac{1}{m}\right\}$, 则 $\{F_m\}$ 是有界闭集列且单调上升到 $G$. 而且 $x \in G^{\mathrm{c}}$ 时 $\rho(x, F_m) \geqslant \dfrac{1}{m}, x \in F_m$ 时 $\rho(x, F_m) = 0$. 所以若取 $\varphi(t)$ 如 (3.3) 式所定义, 并取

$$f_m(x) = \varphi(m\rho(x, F_m)), \tag{4.10}$$

则有

$$\mathbf{1}_G \geqslant f_m \geqslant \mathbf{1}_{F_m}, \quad f_m \in C_{b,b} \quad (m \geqslant 1). \tag{4.11}$$

由于 $\mu_n \overset{\text{lw}}{\Longrightarrow} \mu$, 所以由 (4.11) 有

$$\begin{aligned}\liminf_{n\to\infty} \mu_n(G) &\geqslant \lim_{n\to\infty} \int_{\boldsymbol{B}} f_m d\mu_n \\ &= \int_{\boldsymbol{B}} f_m d\mu \geqslant \mu(F_m). \end{aligned}\tag{4.12}$$

在 (4.12) 中令 $m \to \infty$ 即得 (4.4).

(2) ⇒ (3). 设 (2) 成立. 任取有界 $\mu$ 连续集 $A$, 则

$$\begin{aligned}\mu(A) = \mu(\overline{A}) &\geqslant \lim_{n\to\infty} \sup \mu_n(\overline{A}) \\ &\geqslant \lim_{n\to\infty} \inf \mu_n(A^\circ) \geqslant \mu(A^\circ) = \mu(A),\end{aligned}$$

此即 (3) 成立.

(3) ⇒ (1). 设 (3) 成立. 任取 $f \in C_{b,b}$ 并记 $f$ 的支撑 $\operatorname{supp} f = F$, 则 $F$ 是有界闭集. 对任意 $\varepsilon > 0$, 记

$$F_\varepsilon = \{x \in \boldsymbol{B} : \rho(x, F) \leqslant \varepsilon\},$$

它是有界闭集. 显然 $F_\varepsilon$ 之边界 $\partial F_\varepsilon \subset \{x \in \boldsymbol{B} : \rho(x, F) = \varepsilon\}$. 因此当 $\varepsilon_1 \neq \varepsilon_2$ 时, $\partial F_{\varepsilon_1}$ 与 $\partial F_{\varepsilon_2}$ 不相交. 所以对任何固定的 $\varepsilon > 0$, 在 $(0, \varepsilon)$ 内至多有可数个 $\delta \in (0, \varepsilon)$ 使

$$\mu(\partial F_\delta) > 0,$$

从而总存在 $\varepsilon_0 \in (0,\varepsilon)$, 使 $\mu(\partial F_{\varepsilon_0}) = 0$, 即 $F_{\varepsilon_0}$ 是有界的 $\mu$ 连续集.

设 $A$ 是任意一个 $\mu$ 连续集, 则 $A\bigcap F_{\varepsilon_0}$ 是有界 $\mu$ 连续集. 若令 $\overline{\mu}(B) = \mu(B\bigcap F_{\varepsilon_0}), \overline{\mu}_n(B) = \mu_n(BF_{\varepsilon_0})$, 则 $\overline{\mu}_n(A) = \mu_n(A\bigcap F_{\varepsilon_0}) \to \overline{\mu}(A)$. 所以, 由定理 3.3 知

$$\overline{\mu}_n \stackrel{\mathrm{w}}{\Longrightarrow} \overline{\mu}.$$

于是

$$\int_B f d\mu_n = \int_B f d\overline{\mu}_n \to \int_B f d\overline{\mu} = \int_B f d\mu,$$

此即 $\mu_n \stackrel{\mathrm{lw}}{\Longrightarrow} \mu$. 定理证毕.

**定理 4.2 (局部弱收敛的极限的唯一性)** 设 $\{\mu,\nu,\mu_n,n\geqslant 1\}\subset M_b$, 且

$$\mu_n \stackrel{\mathrm{lw}}{\Longrightarrow} \mu, \quad \mu_n \stackrel{\mathrm{lw}}{\Longrightarrow} \nu, \tag{4.13}$$

则 $\mu$ 与 $\nu$ 在 $\mathscr{B}$ 上恒等.

**证** 任取有界闭集 $F$, 仿照定理 4.1 中证明 “(3) ⇒ (1)” 的办法, 必存在有界闭集列 $\{F_m, m\geqslant 1\}$, 使每个 $F_m$ 既是 $\mu$ 连续集又是 $\nu$ 连续集, 而且 $F_m \downarrow F$ (当 $m\uparrow\infty$ 时). 由 (4.13) 有

$$\nu(F_m) = \lim_{n\to\infty}\mu_n(F_m) = \mu(F_m). \tag{4.14}$$

在 (4.14) 中令 $m\to\infty$ 得

$$\nu(F) = \mu(F) \quad (\text{对任何有界闭集 } F), \tag{4.15}$$

而 $\{\nu,\mu\}\subset M_b$, 用引理 3.1 和 (4.15) 式知: 对任意有界 Borel 集 $A$, 有 $\nu(A) = \mu(A)$. 取闭球列 $S_n \uparrow \mathscr{B}$, 则对任何 $B\in\mathscr{B}$, 有 $\nu(B) = \lim\limits_{n\to\infty}\nu(BS_n) = \lim\limits_{n\to\infty}\mu(BS_n) = \mu(B)$.

下面简单地介绍一下淡收敛.

**定义 4.2** 设 $\{\mu,\mu_n,n\geqslant 1\}\subset M_k$, 若

$$\lim_{n\to\infty}\int_B f d\mu_n = \int_B f d\mu \quad (\forall f\in C_{b,k}),$$

则称 $\{\mu_n\}$ **淡收敛**到 $\mu$, 记之为

$$\lim_{n\to\infty}\mu_n = \mu, \ \text{v.}, \ 或 \ \mu_n \stackrel{\mathrm{v}}{\Longrightarrow} \mu. \tag{4.16}$$

**命题 4.1** 若 $(B,\mathscr{B}(B),\rho)$ 是局部紧可分的完备的距离空间, $\{\mu,\mu_n,n\geqslant 1\}\subset M_b$, 则

$$\mu_n \stackrel{\mathrm{v}}{\Longrightarrow} \mu \Leftrightarrow \mu_n \stackrel{\mathrm{lw}}{\Longrightarrow} \mu. \tag{4.17}$$

**证** 因为局部紧可分的完备的距离空间中的有界闭集都是紧集 (参见 [5] P.8), 所以命题 4.1 成立.

**定理 4.3** 设 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \rho)$ 是局部紧的可分的完备的距离空间, $\{\mu, \mu_n, n \geqslant 1\} \subset M$, 则下列陈述等价:

(1) $\mu_n \stackrel{\text{w}}{\Longrightarrow} \mu$;

(2) $\mu_n(\boldsymbol{B}) \to \mu(\boldsymbol{B})$ 且 $\mu_n \stackrel{\text{lw}}{\Longrightarrow} \mu$;

(3) $\mu_n(\boldsymbol{B}) \to \mu(\boldsymbol{B})$ 且 $\mu_n \stackrel{\text{v}}{\Longrightarrow} \mu$.

**证** (1) $\Rightarrow$ (2) $\Rightarrow$ (3) 显然成立, 再注意命题 4.1, 为证定理 4.3, 只需证明 (2) $\Rightarrow$ (1).

任给 $\varepsilon > 0$, 由于 $\mu(\boldsymbol{B}) < \infty$, 所以由 [4] 和第一章定理 1.4 (关于胎紧的命题) 知: 存在紧集 $K$, 使 $K$ 之补集的 $\mu$ 测度 $\mu(K^{\text{c}}) < \varepsilon$. 仿定理 4.1 的 (3) $\Rightarrow$ (1) 的证明的方法, 可以找到包含 $K$ 的有界闭集 $F_{\varepsilon_0}$, 使 $\mu(\partial F_{\varepsilon_0}) = 0$. 所以, 由

$$\mu_n \stackrel{\text{lw}}{\Longrightarrow} \mu$$

知

$$\mu_n(F_{\varepsilon_0}) \to \mu(F_{\varepsilon_0}). \tag{4.18}$$

而

$$\lim_{n\to\infty} \mu_n(\boldsymbol{B}) = \mu(\boldsymbol{B}) < \infty, \quad \mu(F_{\varepsilon_0}^{\text{c}}) < \varepsilon, \tag{4.19}$$

所以存在正整数 $n_0$, 使

$$\mu_n(F_{\varepsilon_0}^{\text{c}}) < 2\varepsilon (\forall n \geqslant n_0). \tag{4.20}$$

任取 $f \in C_b$, 不妨设 $|f| \leqslant A$ ($A$ 为一正实数). 令

$$\overline{\mu}_n(B) = \mu_n(F_{\varepsilon_0} \textstyle\bigcap B), \quad \overline{\mu}(B) = \mu(F_{\varepsilon_0} \textstyle\bigcap B), \quad \forall B \in \mathscr{B}(\boldsymbol{B}).$$

易证

$$\overline{\mu}_n \stackrel{\text{w}}{\Longrightarrow} \overline{\mu}. \tag{4.21}$$

所以当 $n \geqslant n_0$ 时

$$\begin{aligned}
&\left|\int_{\boldsymbol{B}} f d\mu_n - \int_{\boldsymbol{B}} f d\mu\right| \\
\leqslant &\left|\int_{F_{\varepsilon_0}} f d\mu_n - \int_{F_{\varepsilon_0}} f d\mu\right| + \left|\int_{F_{\varepsilon_0}^{\text{c}}} f d\mu_n - \int_{F_{\varepsilon_0}} f d\mu\right| \\
= &\left|\int_{\boldsymbol{B}} f d\overline{\mu}_n - \int_{\boldsymbol{B}} f d\overline{\mu}\right| + 3A\varepsilon.
\end{aligned} \tag{4.22}$$

由 $\overline{\mu}_n \stackrel{\text{w}}{\Longrightarrow} \overline{\mu}$ 及 $\varepsilon > 0$ 可任意小, 在 (4.22) 中先令 $n \to \infty$ 再令 $\varepsilon \to 0$ 可知 $\mu_n \stackrel{\text{w}}{\Longrightarrow} \mu$. 定理 4.3 得证.

**附注 4.1**　弱收敛、局部弱收敛和淡收敛是三个不同的概念.

**例 4.1**　取 $\boldsymbol{B}=\{1,2,3,\cdots\}$, 在 $\boldsymbol{B}$ 中定义距离 $\rho(i,j)=\delta_{i,j}=0$ 或 1 当 $i\neq j$ 或 $i=j$.$\boldsymbol{B}$ 中之拓扑为离散拓扑, 即 $\boldsymbol{B}$ 中任一子集既是开集又是闭集. 显然 $\boldsymbol{B}$ 中的子集 $K$ 是紧集的充分必要条件是: $K$ 是有限集. 又 $\operatorname{diam}(\boldsymbol{B})=1$, 所以 $(\boldsymbol{B},\rho)$ 是有界距离空间, 故 $\mathscr{B}(\boldsymbol{B})$ 上的测度列的弱收敛与局部弱收敛的概念是一致的. 但局部弱收敛的概念与淡收敛的概念是不一致的. 例如, 取 $\mu_n(\{i\})=\delta_{n,i}$, 取 $\mu(\{i\})\equiv 0$. 任取紧集 $K,K$ 必为有限集, 且是 $\mu$ 连续集, 故 $\lim\limits_{n\to\infty}\mu_n(K)=0=\mu(K)$, 从而 $\mu_n\overset{\mathrm{v}}{\Longrightarrow}\mu$. 但 $\boldsymbol{B}$ 是 $\mu$ 的有界连续集, 而 $\lim\limits_{n\to\infty}\mu_n(\boldsymbol{B})=1\neq 0=\mu(\boldsymbol{B})$, 这说明 $\{\mu_n\}$ 不局部弱收敛到 $\mu$.

**例 4.2**　取 $\boldsymbol{B}=\mathbf{R},\rho$ 是 $\mathbf{R}$ 中的欧氏距离. 令 $\mu_n(\{n\})=1,\mu_n(\mathbf{R}-\{n\})=0,\mu\equiv 0$, 则对任何有界集 $A$, 有 $\mu_n(A)\to 0=\mu(A)$, 故 $\mu_n\overset{\mathrm{lw}}{\Longrightarrow}\mu$. 但 $\lim\limits_{n\to\infty}\mu_n(\mathbf{R})=1\neq 0=\mu(\boldsymbol{B})$, 所以 $\{\mu_n\}$ 不弱收敛于 $\mu$.

## §5　欧氏空间中的特殊场合

在第 3 节和第 4 节中, 我们研究了一般的距离空间中的测度列的弱收敛、局部弱收敛和淡收敛及它们的关系. 现在我们把距离空间 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$ 特殊化为 $(\mathbf{R}^N,\mathscr{B}(\mathbf{R}^N),\rho_N)$, 其中 $\rho_N$ 为 $\mathbf{R}^N$ 中的欧氏距离, 即是对任何 $x=(x_1,\cdots,x_N)\in\mathbf{R}^N,y=(y_1,\cdots,y_N)\in\mathbf{R}^N$, 定义

$$\rho_N(x,y)=\sqrt{\sum_{i=1}^{N}(x_i-y_i)^2}.$$

众所周知 $(\mathbf{R}^N,\mathscr{B}(\mathbf{R}^N),\rho_N)$ 是局部紧的可分的完备的距离空间. 所以, 由命题 4.1, 在欧氏空间的特殊场合, 淡收敛可以不必讨论. 下面只讨论弱收敛与局部弱收敛. 沿用第一章 §4 的符号.

**定义 5.1**　设 $F(x)$ 和 $F_n(x)$ 都是测度有界的 (L-S)$_1$ 函数 (定义见第一章定义 4.1) $(n\geqslant 1)$. 如果对 $F(x)$ 的任一连续点 $x_0$ (即 $F(x_0+0)-F(x_0-0)=0$) 都有

$$\lim_{n\to\infty}F_n(x_0)=F(x_0),\tag{5.1}$$

则称 $\{F_n(x),n\geqslant 1\}$ 局部弱收敛到 $F(x)$, 记作

$$\lim_{n\to\infty}F_n(x)=F(x),\ \text{lw.},\quad 或\quad F_n(x)\xrightarrow{\mathrm{lw}}F(x),\quad 或\quad F_n\xrightarrow{\mathrm{lw}}F.$$

如果 $F_n\xrightarrow{\mathrm{lw}}F$ 且 $F_n(\infty)\to F(\infty),F_n(-\infty)\to F(-\infty)$, 则称 $\{F_n(x),n\geqslant 1\}$ 弱

收敛到 $F(x)$, 记作

$$\lim_{n\to\infty} F_n(x) = F(x), \quad \text{w.}, \quad 或 \quad F_n(x) \xrightarrow{\text{w}} F(x), \quad 或 \quad F_n \xrightarrow{\text{w}} F.$$

$C(F)$ 表示 $F(x)$ 的全部连续点集.

**命题 5.1** (1) 若 $\{F(x), F_n(x), n \geqslant 1\}$ 是测度有界的 $(\text{L-S})_1$ 函数列, 且 $F_n(x) \xrightarrow{\text{lw}} F(x)$ (相应地, $F_n(x) \xrightarrow{\text{w}} F(x)$), 则 $\{F(x), F_n(x) n \geqslant 1\}$ 的对应的 $(\text{L-S})_1$ 测度列 $\{\mu_F, \mu_{F_n}, n \geqslant 1\}$ 满足 $\mu_{F_n} \overset{\text{lw}}{\Longrightarrow} \mu_F$ (相应地, $\mu_{F_n} \overset{\text{w}}{\Longrightarrow} \mu_F$).

(2) 若 $\{\mu, \mu_n, n \geqslant 1\} \subset M, \mu_n \overset{\text{lw}}{\Longrightarrow} \mu$ (相应地, $\mu_n \overset{\text{w}}{\Longrightarrow} \mu$), 则存在测度有界的 $(\text{L-S})_1$ 函数列 $\{F(x), F_n(x), n \geqslant 1\}$, 使

$$\mu = \mu_F, \quad \mu_n = \mu_{F_n} \quad (n \geqslant 1),$$

且 $F_n(x) \xrightarrow{\text{lw}} F(x)$ (相应地, $F_n(x) \xrightarrow{\text{w}} F(x)$).

**证** (1) 若 $\{F(x), F_n(x), n \geqslant 1\}$ 是测度有界的 $(\text{L-S})_1$ 函数列, 令 $\{\mu_F, \mu_{F_n}, n \geqslant 1\}$ 为其相应的 $(\text{L-S})_1$ 测度, 则 $\{\mu_F, \mu_{F_n}, n \geqslant 1\} \subset M$. 由 $F_n \xrightarrow{\text{lw}} F$ 知: 对 $\mu_F$ 的任何连续区间 $(a, b], a$ 和 $b$ 都是 $F(x)$ 的连续点, 所以

$$\mu_{F_n}((a, b]) = F_n(b) - F_n(a) \to F(b) - F(a) = \mu_F((a, b]).$$

由此易证: 对 $\mu_F$ 的任何有界连续集 $A$, 都有

$$\mu_{F_n}(A) \to \mu_F(A).$$

所以 $\mu_{F_n} \overset{\text{lw}}{\Longrightarrow} \mu_F$.

仿之可证: 当 $F_n \xrightarrow{\text{w}} F$ 时, 有 $\mu_{F_n} \overset{\text{w}}{\Longrightarrow} \mu_F$.

(2) 若 $\{\mu, \mu_n, n \geqslant 1\} \subset M$, 令

$$F(x) = \begin{cases} \mu((0, x]), & 当\ x > 0, \\ 0, & 当\ x = 0, \\ -\mu((x, 0]), & 当\ x < 0. \end{cases}$$

$F_n(x)$ 亦仿上定义 (用 $\mu_n$ 取代 $\mu$), 则 $\{F(x), F_n(x), n \geqslant 1\}$ 即为所求.

**定理 5.1** 设 $\{\mu, \mu_n, n \geqslant 1\}$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的一列测度, 且 $\mu(I) < \infty, \mu_n(I) < \infty, (\forall n \geqslant 1, I \in \mathscr{I}_N)$. 若对每个 $\mu$ 连续区间 $I$, 都有 $\mu_n(I) \to \mu(I)$, 则

$$\mu_n \overset{\text{lw}}{\Longrightarrow} \mu.$$

**证** 对每个开集 $G$, 均可表为至多可数个互不相交的区间之并, 而每个区间又可表为至多可数个互不相交的 $\mu$ 连续区间之并. 所以

$$G = \bigcup_{i \geqslant 1} I_i,$$

$\{I_i\}$ 是两两不交的 $\mu$ 连续区间列. 所以

$$\mu(G) = \lim_{k\to\infty} \mu\left(\bigcup_{i=1}^{k} I_i\right) = \lim_{k\to\infty} \lim_{n\to\infty} \mu_n\left(\bigcup_{i=1}^{k} I_i\right) \leqslant \lim_{n\to\infty} \inf \mu_n(G).$$

对每个有界闭集 $F$, 必存在区间 $I \in \mathscr{I}_N$, 使 $I^\circ \supset F$ 且 $I$ 是 $\mu, \mu_1, \mu_2, \cdots$ 的连续区间. 所以

$$\mu(I) = \lim_{n\to\infty} \mu_n(I) = \lim_{n\to\infty} \mu_n(I^\circ) = \mu(I^\circ).$$

而 $I^\circ - F$ 是开集, 所以

$$\mu(I^\circ - F) \leqslant \lim_{n\to\infty} \inf \mu_n(I^\circ - F).$$

因此,

$$\mu(F) \geqslant \lim_{n\to\infty} \sup[\mu_n(I^\circ) - \mu_n(I^\circ - F)] \geqslant \lim_{n\to\infty} \sup \mu_n(F).$$

所以由定理 4.1 知 $\mu_n \xrightarrow{\text{lw}} \mu$.

**定理 5.2** 设 $\{\mu, \mu_n, n \geqslant 1\}$ 如定理 5.1, 则下述陈述等价:

(1) 对每个 $\mu$ 连续区间 $I \in \mathscr{I}_N$, 有 $\mu_n(I) \to \mu(I)$;

(2) $\mu_n \xrightarrow{\text{lw}} \mu$;

(3) 对任何开集 $G, \mu(G) \leqslant \lim\limits_{n\to\infty} \inf \mu_n(G)$; 对任何有界闭集 $F, \mu(F) \geqslant \lim\limits_{n\to\infty} \sup \mu_n(F)$;

(4) 对任何 $\mu$ 有界连续集 $A, \mu_n(A) \to \mu(A)$

**证** 由定理 4.1 和定理 5.1 即得定理 5.2.

**定理 5.3** 设 $\{\mu, \mu_n, n \geqslant 1\}$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的一列测度, 且 $\mu(\mathbf{R}^N) < \infty$, $\mu_n(\mathbf{R}^N) < \infty (\forall n \geqslant 1)$, 则下述陈述等价:

(1) 对每个 $\mu$ 连续区间 $I \in \mathscr{I}_N$, 有 $\mu_n(I) \to \mu(I)$, 且 $\mu_n(\mathbf{R}^N) \to \mu(\mathbf{R}^N)$;

(2) $\mu_n \xrightarrow{\text{lw}} \mu$, 且 $\mu_n(\mathbf{R}^N) \to \mu(\mathbf{R}^N)$

(3) $\mu_n \xrightarrow{\text{w}} \mu$.

**证** 由定理 4.3 和定理 5.2 即得定理 5.3.

**定义 5.2**　设 $\mu$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的一个测度, 记

$$D(\mu)=\left\{a\in\mathbf{R}:\begin{array}{l}\text{存在一个 } 1\leqslant i\leqslant N,\text{ 使}\\ \mu(\{x\in\mathbf{R}^N:x_i=a\})>0\end{array}\right\},$$
$$c(\mu)=\mathbf{R}-D(\mu).$$

注意: $\{x\in\mathbf{R}^N:x_i=a\}$ 是 $\mathbf{R}^N$ 空间中的 $N-1$ 维超平面.

如果 $I\in\mathscr{I}_N$ 是一个 $N$ 维区间, 而且它的每一个顶点的每一个坐标都在 $\mathbf{R}$ 中某子集 $D$ 中, 则称 $I$ 是一个 $D$ **区间**.

显然, 若 $\mu$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的一个测度, 每个 $C(\mu)$ 区间必是 $\mu$ 的连续区间. 当 $\mu$ 在 $\mathscr{I}_N$ 有限时, $C(\mu)$ 是 $\mathbf{R}$ 中的处处稠密子集.

**定义 5.3**　设 $D$ 是 $\mathbf{R}$ 中的稠密子集, $\mu$ 是定义在 $\mathscr{I}_N$ 中的一切 $D$ 区间上的集合函数, 满足:

(1) $\mu(\varnothing)=0$ 且 $0\leqslant\mu(I)<\infty$ (对一切 $D$ 区间 $I$);

(2) $\mu$ 在全体 $D$ 区间上有有限可加性, 则称 $\mu$ 是 $D_N$ **次测度**.

**定理 5.4**　设 $D$ 是 $\mathbf{R}$ 中稠密子集, $\mu$ 是 $D_N$ 次测度, 则在 $\mathscr{B}(\mathbf{R}^N)$ 上存在唯一一个测度 $\mu^*$, 使得对任意 $D$ 区间 $I$ 和区间 $J\in\mathscr{I}_N$,

$$\text{(1)}\qquad I^{\mathrm{c}}\subset J^{\circ}\Rightarrow\mu(I)\leqslant\mu^*(J);\tag{5.2}$$

$$\text{(2)}\qquad I^{\mathrm{c}}\supset J^{\mathrm{c}}\Rightarrow\mu(I)\geqslant\mu^*(J).\tag{5.3}$$

称 $\mu^*$ 为 $\mu$ 的伴随测度.

**证**　任取 $I=(a,b]\in\mathscr{I}_N$, 定义

$$\mu^*(I)=\mu^*((a,b])=\lim_{\substack{a'\downarrow a\\ b'\downarrow b}}\mu((a',b']),\tag{5.4}$$

其中 $(a',b']$ 是 $D$ 区间. 可证 $\mu^*$ 是 $\mathscr{I}_N$ 上的有限测度.

事实上 $\mu^*$ 满足:

(i) $\mu^*(\varnothing)=0$, 且 $0\leqslant\mu^*(I)<\infty(\forall I\in\mathscr{I}_N)$;

(ii) $\mu^*$ 在 $\mathscr{I}_N$ 有有限可加性;

(iii) 对任何 $a=(a_1,\cdots,a_N),b=(b_1,\cdots,b_N)$, 任何 $1\leqslant i\leqslant N$, 有

$$\lim_{b_i\downarrow a_i}\mu^*((a,b])=0.$$

由上述三性质易证 $\mu^*$ 在 $\mathscr{I}_N$ 上有可数可加性, 所以 $\mu^*$ 是 $\mathscr{I}_N$ 上的一个有限测度, 从而它可以唯一地扩张到 $\sigma(\mathscr{I}_N)=\mathscr{B}(\mathbf{R}^N)$ 上去而得一个 $\sigma$ 有限测度, 记

此扩张后的测度仍为 $\mu^*$, 由 $\mu^*$ 的定义可验证 (5.2) 和 (5.3) 成立. 于是伴随测度的存在性获证.

若还有另一个伴随测度 $\mu_1^*$. 任给 $\varepsilon > 0$, 任取 $J \in \mathscr{I}_N$, 则存在 $J' \in \mathscr{I}_N$, 使

$$(J')^\circ \supset J^c, \quad \mu^*(J) \geqslant \mu^*(J'') - \varepsilon. \tag{5.5}$$

再取 $D$ 区间 $I$, 使

$$I^c \subset (J')^\circ, \quad I^\circ \supset J^c,$$

故

$$\mu^*(J') \geqslant \mu(I) \geqslant \mu_1^*(J). \tag{5.6}$$

由 (5.5) 和 (5.6) 得:

$$\mu^*(J) \geqslant \mu_1^*(J) - \varepsilon,$$

由 $\varepsilon > 0$ 可任意小知

$$\mu^*(J) \geqslant \mu_1^*(J) \quad (\forall J \in \mathscr{I}_N). \tag{5.7}$$

由 $\mu^*$ 与 $\mu_1^*$ 地位的对称性得

$$\mu_1^*(J) \geqslant \mu^*(J) \quad (\forall J \in \mathscr{I}_N). \tag{5.8}$$

由 (5.7), (5.8) 得 $\mu^* \equiv \mu_1^*$. 定理证毕.

**定理 5.5** *$\mathscr{B}(\mathbf{R}^N)$ 上的在 $\mathscr{I}_N$ 上有限的测度列 $\{\mu_n, n \geqslant 1\}$ 局部弱收敛的充分必要条件是: 存在一个在 $\mathbf{R}$ 中处处稠密的子集 $D$, 使得对任何 $D$ 区间 $I$, 都有*

$$\lim_{n\to\infty} \mu_n(I) \quad \text{存在 (为有限实数)}.$$

**证** 必要性. 若 $\mu_n \xrightarrow{\text{lw}} \mu$, 则对任何 $c(\mu)$ 区间 $I$ (它必是 $\mu$ 的连续区间), 有 $\lim\limits_{n\to\infty} \mu_n(I) = \mu(I)$ 存在 (为有限数). 而 $c(\mu)$ 在 $\mathbf{R}$ 中处处稠密, 必要性得证.

充分性. 在 $D$ 区间上造一个集合函数 $\mu$ 如下:

$$\mu(I) = \lim_{n\to\infty} \mu_n(I), \quad I \text{ 是 } D \text{ 区间},$$

则

(1) $0 \leqslant \mu(I) < \infty$ ($I$ 是任意 $D$ 区间);

(2) $\mu$ 在全体 $D$ 区间上有有限可加性, 故 $\mu$ 是 $D_N$ 次测度. 由定理 5.4 知 $\mu$ 有唯一的一个伴随测度 $\mu^*$ 满足 (5.2) 及 (5.3) 式. 往证

$$\mu_n \xrightarrow{\text{lw}} \mu^*.$$

事实上, 任给一个区间 $I \in \mathscr{I}_N$, 存在一列 $D$ 区间 $\{I_l\}$, 使 $I_l^\circ \supset I_{l+1}^c, I_l^\circ \supset I, (l=1,2,\cdots), \lim\limits_{l\to\infty} I_l = I^c$. 由于 $\mu^*$ 是 $\mu$ 的伴随测度, 又因为 $I_l^\circ \supset I_{l+1}^c$, 所以

$$\mu^*(I_l) \geqslant \mu(I_{l+1}).$$

但是 $I_{l+1}$ 是 $D$ 区间, 所以

$$\lim_{n\to\infty} \mu_n(I_{l+1}) = \mu(I_{l+1}).$$

而 $I_{l+1} \supset I$, 所以

$$\lim_{n\to\infty} \mu_n(I_{l+1}) \geqslant \limsup_{n\to\infty} \mu_n(I).$$

因此

$$\mu^*(I^c) = \lim_{l\to\infty} \mu^*(I_l) \geqslant \limsup_{n\to\infty} \mu_n(I).$$

仿之可证

$$\mu^*(I^\circ) \leqslant \liminf_{n\to\infty} \mu_n(I).$$

如果 $I$ 是 $\mu^*$ 的连续区间, 则由上述两个不等式有:

$$\lim_{n\to\infty} \mu_n(I) = \mu^*(I).$$

此即 $\mu_n \xrightarrow{\text{lw}} \mu^*$. 定理证毕.

**定理 5.6** 设 $\{\mu_n, n \geqslant 1\}$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的测度列, 且对任何 $I \in \mathscr{I}_N, \{\mu_n(I), n \geqslant 1\}$ 有界 $G(I)$, 则 $\{\mu_n, n \geqslant 1\}$ 有局部弱收敛的子列. (此性质通常称之为 $\{\mu_n\}$ 的局部弱紧性.)

**证** 由定理 5.5, 为证定理 5.6, 只需证明 $\{\mu_n\}$ 中有一个子列, 它在一切 $D$ 区间上收敛, 而 $D$ 是 $\mathbf{R}$ 中的一个稠密子集.

取 $D$ 为有理数集, 则 $D$ 区间总共有可数个, 记之为 $I_1, I_2, \cdots$.

由于 $\{\mu_n(I_1), n \geqslant 1\}$ 有界 $G(I_1)$, 所以它有收敛子列 $\{\mu_{n,1}(I_1), n \geqslant 1\}$. 又由于 $\{\mu_{n,1}(I_2), n \geqslant 1\}$ 有界 $G(I_2)$, 所以它又有收敛子列 $\{\mu_{n,2}(I_2), n \geqslant 1\}, \cdots$, 如此继续下去, 对每个正整数 $k, \{\mu_n, n \geqslant 1\}$ 都有子列 $\{\mu_{n,k}, n \geqslant 1\}$, 它在 $I_1, I_2, \cdots, I_k$ 上均收敛. 取 $\{\mu_{n,n}, n \geqslant 1\}$, 它必在 $I_1, I_2, I_3, \cdots$ 上均收敛, 即它在一切 $D$ 区间上都收敛. 定理证毕.

**推论 5.1** 设 $\{\mu_n, n \geqslant 1\}$ 满足定理 5.6 中一切条件, 若它不局部弱收敛, 则它必有两个局部弱收敛子序列, 它们局部弱收敛到不同的极限; 若 $\{\mu_n, n \geqslant 1\}$ 的任一局部弱收敛的子列都局部弱收敛到同一极限 $\mu$, 则 $\{\mu_n, n \geqslant 1\}$ 也局部弱收敛到 $\mu$.

**证** 由定理 5.6, $\{\mu_n\}$ 必有局部弱收敛子列, 记之为 $\{\mu_{k,k}, k \geqslant 1\}, \mu_{k,k} \xrightarrow{\text{lw}} \mu$. $\{\mu_n\}$ 抽去子序列 $\{\mu_{k,k}, k \geqslant 1\}$ 后所剩下的子列记之为 $\{\mu_{n'}\}$. 若 $\{\mu_n\}$ 不局部弱收敛, 则 $\{\mu_{n'}\}$ 不局部弱收敛到 $\mu$. 所以存在一个 $\mu$ 连续间 $I$, 使实数列 $\{\mu_{n'}(I)\}$ 不收敛到 $\mu(I)$. 因此 $\{\mu_{n'}(I)\}$ 有子列 $\{\mu_{k',k'}(I)\}$, 使

$$|\mu_{k',k'}(I) - \mu(I)| > \varepsilon \quad (\forall k' \geqslant 1). \tag{5.9}$$

再一次应用定理 5.6, 得知 $\{\mu_{k',k'}\}$ 有局部弱收敛子列 $\{\mu_{k_j}, j \geqslant 1\}$ 使

$$\mu_{k_j} \xrightarrow{\text{lw}} \nu, \quad (j \to \infty).$$

由 (5.9) $\mu$ 与 $\nu$ 在 $\mathscr{B}(\mathbf{R}^N)$ 上是不能恒等的. 推论 5.1 证毕.

**定理 5.7** 设 $\{\mu, \mu_n, n \geqslant 1\}$ 是 $\mathscr{B}(\mathbf{R}^N)$ 上的测度列且 $\mu(\mathbf{R}^N) \leqslant K, \mu_n(\mathbf{R}^N) \leqslant K \mu_n \xrightarrow{\text{lw}} \mu$, 则 $\mu_n \xrightarrow{\text{w}} \mu$ 的充分必要条件是: 任给 $\varepsilon > 0$, 存在正整数 $n_0$ 及区间 $I_0 \in \mathscr{I}_N$ 使得

$$\mu_n(\mathbf{R}^N) - \mu_n(I) < \varepsilon \quad (\forall n \geqslant n_0, I \supset I_0). \tag{5.10}$$

**证** 必要性. 任给 $\varepsilon > 0$, 对 $\mu$ 而言, 存在一个 $\mu$ 连续区间 $I_0 \in \mathscr{I}_N$, 使任何区间 $I \supset I_0$ 必有

$$\mu(\mathbf{R}^N) - \mu(I) < \frac{\varepsilon}{2}. \tag{5.11}$$

对 $\dfrac{\varepsilon}{2} > 0$, 又存在一个正整数 $n_0$, 使 $n \geqslant n_0$ 时有

$$|\mu_n(\mathbf{R}^N) - \mu(\mathbf{R}^N)| < \frac{\varepsilon}{4}, \quad |\mu_n(I_0) - \mu(I_0)| < \frac{\varepsilon}{4}. \tag{5.12}$$

所以, 当 $n \geqslant n_0$, 且 $I \supset I_0$ 时有

$$\begin{aligned}\mu_n(\mathbf{R}^N) - \mu_n(I) &\leqslant \mu_n(\mathbf{R}^N) - \mu_n(I_0) \\ &\leqslant |\mu_n(\mathbf{R}^N) - \mu(\mathbf{R}^N)| + |\mu(\mathbf{R}^N) - \mu(I_0)| \\ &\quad + |\mu(I_0) - \mu_n(I_0)| < \varepsilon.\end{aligned}$$

充分性. 设对任何 $\varepsilon > 0$, 存在 $I_0 \in \mathscr{I}_N$ 及正整数 $n_0$ 使 (5.10) 成立. 不妨令 $I_0$ 是 $\mu$ 的连续区间, 且 $|\mu(I_0) - \mu(\mathbf{R}^N)| < \varepsilon$, 于是

$$\begin{aligned}&|\mu_n(\mathbf{R}^N) - \mu(\mathbf{R}^N)| \leqslant |\mu_n(\mathbf{R}^N) - \mu_n(I_0)| \\ &+ |\mu_n(I_0) - \mu(I_0)| + |\mu(I_0) - \mu(\mathbf{R}^N)| \\ &\leqslant 2\varepsilon + |\mu_n(I_0) - \mu(I_0)| \quad (\forall n \geqslant n_0).\end{aligned}$$

在上式中令 $n \to \infty$ 并注意 $\mu_n \xrightarrow{\text{lw}} \mu$ 得:

$$\limsup_{n\to\infty} |\mu_n(\mathbf{R}^N) - \mu(\mathbf{R}^N)| \leqslant 2\varepsilon.$$

由 $\varepsilon > 0$ 可以任意小知

$$\lim_{n\to\infty} \mu_n(\mathbf{R}^N) = \mu(\mathbf{R}^N)$$

所以 $\mu_n \xrightarrow{\text{w}} \mu$. 定理证毕.

**附注 5.1** (1) 确有测度有界的 $(\text{L-S})_1$ 函数列 $\{F_n(x), n \geqslant 1\}$, 它所产生的 $(\text{L-S})_1$ 测度列 $\{\mu_{F_n}, n \geqslant 1\}$ 局部弱收敛到某个测度有界的 $(\text{L-S})_1$ 函数 $F(x)$ 所产生的 $(\text{L-S})_1$ 测度 $\mu_F$, 但 $\{F_n(x), n \geqslant 1\}$ 并不局部弱收敛到 $F(x)$. 反例如下.

取 $F(x) \equiv 0, \mu_F \equiv 0$,

$$F_0(x) = \begin{cases} 0, & \text{当 } x < 0, \\ 1, & \text{当 } x \geqslant 0, \end{cases} \quad F_n(x) = F_0(x+n), (n \geqslant 1),$$

则

$$\lim_{n\to\infty} \mu_{F_n}((a,b]) \equiv 0, \quad (\forall a, b \in \mathbf{R}).$$

所以

$$\mu_{F_n} \xrightarrow{\text{lw}} \mu_F.$$

但是

$$\lim_{n\to\infty} F_n(x) \equiv 1 \quad (\forall x \in \mathbf{R}),$$

所以 $\{F_n(x), n \geqslant 1\}$ 绝不可能局部弱收敛到 $F(x)$.

(2) 局部弱收敛的测度列而不弱收敛的测度列是存在的. 例如 (1) 中的 $\{\mu_{F_n}, n \geqslant 1\}$, 有

$$\mu_{F_n} \xrightarrow{\text{lw}} \mu_F,$$

但是

$$\mu_{F_n}(\mathbf{R}^1) \equiv 1 \quad (\forall n \geqslant 1), \quad \mu_F(\mathbf{R}^1) = 0,$$

所以 $\{\mu_{F_n}, n \geqslant 1\}$ 绝不可能弱收敛到 $\mu_F$.

(3) 当 $\mu_n \xrightarrow{\text{lw}} \mu$ 时, 确有开集 $G$ 和闭集 $K$, 使

$$\lim_{n\to\infty} \mu_n(G) \neq \mu(G), \quad \lim_{n\to\infty} \mu_n(K) \neq \mu(K).$$

反例如下:

取

$$F_1(x)=\begin{cases}0, & \text{当 } x<0,\\ 1, & \text{当 } x\geqslant 0,\end{cases}$$

$$F_n(x)=F_1\left(x+1-\frac{1}{n}\right)$$

$$=\begin{cases}0, & \text{当 } x<-1+\dfrac{1}{n},\\ 1, & \text{当 } x\geqslant -1+\dfrac{1}{n},\end{cases}\qquad n=2,3,\cdots,$$

$$F(x)=\begin{cases}0, & \text{当 } x<-1,\\ 1, & \text{当 } x\geqslant -1,\end{cases}$$

则

$$\mu_{F_n}\overset{\text{lw}}{\Longrightarrow}\mu_F.$$

但是

$$\begin{aligned}\mu_{F_n}((-1,\infty))&=\lim_{x\to\infty}\mu_{F_n}((-1,x])\\&=\lim_{x\to\infty}F_n(x)-F_n(-1)=1\quad(\forall n\geqslant 2),\\ \mu_F((-1,\infty))&=\lim_{x\to\infty}F(x)-F(-1)=0,\\ \mu_{F_n}([-2,-1])&=F_n(-1)-F_n(-2-0)=0\quad(\forall n\geqslant 2),\\ \mu_F([-2,-1])&=1.\end{aligned}$$

取 $G=(-1,\infty),K=[-2,-1]$ 即为所求.

**定理 5.8** 设 $\{\mu_F,\mu_{F_n},n\geqslant 1\}$ 为 $\mathscr{B}(\mathbf{R}^N)$ 上一列 $(\text{L-S})_N$ 测度且在每个区间 $I\in\mathscr{I}_N$ 上的测度值有限, $\mu_{F_n}\overset{\text{lw}}{\Longrightarrow}\mu_F$. $U$ 为一个抽象集合. $g(u,x);Ux\mathbf{R}^N\mapsto\mathbf{R}$, 且满足条件: 对任何 $\varepsilon>0$ 及区间 $J\in\mathscr{I}_N$, 都存在一个 $\delta>0$, 当 $x,y\in J,\rho_N(x,y)<\delta$ 时, 对一切 $u\in U$, 一致地有

$$|g(u,x)-g(u,y)|<\delta,$$

则对 $\mu_F$ 的任何一个连续区间 $I$, 对 $u\in U$ 一致地有

$$\lim_{n\to\infty}\int_I g(u,x)\mu_{F_n}(dx)=\int_I g(u,x)\mu_F(dx).\tag{5.13}$$

证明甚易, 读者可作习题验证之.

**定理 5.9** 在定理 5.8 的条件下, 若还有 $\mu_{F_n}(\mathbf{R}^N)\leqslant K$ ($K$ 为常数), $\mu_{F_n}(\mathbf{R}^N)\to\mu_F(\mathbf{R}^N),|g(u,x)|\leqslant K(\forall u\in U,x\in\mathbf{R}^N)$, 则对 $u\in U$ 一致地有

$$\lim_{n\to\infty}\int_{\mathbf{R}^N}g(u,x)\mu_{F_n}(dx)=\int_{\mathbf{R}^N}g(u,x)\mu_F(dx).\tag{5.14}$$

证明甚易, 读者可做为习题验证之.

**定理 5.10** 在定理 5.8 的条件下. 若还有

$$\lim_{|x|\to\infty} g(u,x)=0 \quad (\text{对 } u\in U \text{ 一致成立}),$$

且 $\mu_{F_n}(\mathbf{R}^N)\leqslant K, \mu_F(\mathbf{R}^N)\leqslant K$, 则对 $u\in U$, 一致地有:

$$\lim_{n\to\infty}\int_{\mathbf{R}^N} g(u,x)\mu_{F_n}(dx)=\int_{\mathbf{R}^N} g(u,x)\mu_F(dx).$$

证明甚易, 读者可作为习题验证之.

**附注 5.2** 在有的著作中, 例如在文献 [36] 中, 称本书的局部弱收敛 ($\mu_n \xRightarrow{\text{lw}} \mu$) 为弱收敛 (记作 $\mu_n \xrightarrow{\text{w}} \mu$); 而称本书中的弱收敛 ($\mu_n \xRightarrow{\text{w}} \mu$) 为全收敛 (记作 $\mu_n \xrightarrow{\text{c}} \mu$).

## §6 习题及应用

1. 设 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$ 是可测距离空间, $\mathscr{B}'\subset\mathscr{B}(\boldsymbol{B})$. 若 $\mathscr{B}'$ 满足:

(i) 对有限交运算封闭;

(ii) $\boldsymbol{B}$ 中每个开集都可表为 $\mathscr{B}'$ 中有限或可数无穷多个元素之并.

而且概率测度列 $\{P,P_n,n\geqslant 1\}$ 在 $\mathscr{B}'$ 中每个 $A$ 都有 $P_n(A)\to P(A)$. 试证: $P_n \xRightarrow{\text{w}} P$.

2. (续上) 若 $\mathscr{B}'$ 满足

(i) 对有限交封闭;

(ii) 对每个 $x\in\boldsymbol{B},\varepsilon>0$, 存在 $A\in\mathscr{B}'$ 及一个以 $x$ 为中心、以 $\varepsilon$ 为半径的开球 $S(x,\varepsilon)$, 使 $x\in A^\circ\subset A\subset S(x,\varepsilon)$. 若 $\boldsymbol{B}$ 是可分的且对每个 $A\in\mathscr{B}'$, 有 $P_n(A)\to P(A)$, 则 $P_n \xRightarrow{\text{w}} P$.

3. (续上) 若 $\boldsymbol{B}$ 是可分的, 且对每个可表为有限个开球的交的 $P$ 连续集 $A$, 都有 $P_n(A)\to P(A)$, 则 $P_n \xRightarrow{\text{w}} P$.

4. 设 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$ 是离散的距离空间, 即 $\boldsymbol{B}$ 是可数集, $\mathscr{B}(\boldsymbol{B})$ 是 $\boldsymbol{B}$ 中一切子集所构成的 $\sigma$ 代数. $\{P,P_n,n\geqslant 1\}$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的概率测度列, 则 $P_n \xRightarrow{\text{w}} P$ 的充分必要条件是: 对每个 $x\in\boldsymbol{B}$, 都有 $P_n(\{x\})\to P(\{x\})$.

5. 设 $\lambda$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的一个测度, 概率测度列 $\{P,P_n,n\geqslant 1\}$ 可表为

$$P(A)=\int_A p(x)\lambda(dx),\quad P_n(A)=\int_A p_n(x)\lambda(dx),\quad A\in\mathscr{B}(\boldsymbol{B}),$$
$$p_n(x)\to p(x),\quad \lambda-\text{a.s.},$$

则 $P_n \xRightarrow{\text{w}} P$.

6. (乘积空间上的弱收敛性) 设 $(\boldsymbol{B}_i, \mathscr{B}(\boldsymbol{B}_i), \rho_i)$ 都是可测距离空间, 且 $B_i$ 是可分的, $i=1,2$. 令 $\boldsymbol{B}=\boldsymbol{B}_1\times\boldsymbol{B}_2, \mathscr{B}(\boldsymbol{B})=\mathscr{B}(\boldsymbol{B}_1)\times\mathscr{B}(\boldsymbol{B}_2), \rho((x_1,x_2),(y_1,y_2))=\max\{\rho_1(x_1,y_1),\rho_2(x_2,y_2)\}, (x_1,x_2),(y_1,y_2)\in\boldsymbol{B}_1\times\boldsymbol{B}_2$, 则 $(\boldsymbol{B},\mathscr{B}(\boldsymbol{B}),\rho)$ 是可分的可测距离空间. 设 $\{P,P_n,n\geqslant 1\}$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的一列概率测度, 令 $P'(A)=P(A\times\boldsymbol{B}_2), P''(B)=P(\boldsymbol{B}_1\times B), (A\in\boldsymbol{B}_1, B\in\boldsymbol{B}_2)$ 为 $P$ 的边缘分布. $P_n'$ 与 $P_n''$ 亦类似地定义, 则 $P_n \xrightarrow{\text{w}} P$ 的充分必要条件是:

$$P_n(A'\times A'')\to P(A'\times A''),$$

其中 $A'$ 是 $P'$ 连续集, $A''$ 是 $P''$ 连续集.

7. 设 $\{X_n,n\geqslant 1\}$ 和 $\{Y_n,n\geqslant 1\}$ 都是概率空间 $(\Omega,\mathscr{F},P)$ 上的实值随机变量列. $a$ 是常数, $F(x)$ 是标准的 (L-S) 函数 (分布函数). 若

$$P(X_n\leqslant x)\xrightarrow{\text{w}}F(x),\quad Y_n\xrightarrow{\text{P}}a,$$

则

(1) $P(X_n+Y_n\leqslant x)\xrightarrow{\text{w}}F(x-a)$;

(2) $P(X_nY_n\leqslant x)\xrightarrow{\text{w}}F\left(\dfrac{x}{a}\right)$, (当 $a>0$);

(3) $P\left(\dfrac{X_n}{Y_n}\leqslant x\right)\xrightarrow{\text{w}}F(ax)$, (当 $a>0$).

8. 证明定理 5.8.

9. 证明定理 5.9.

10. 证明定理 5.10.

# 第四章　特征函数和特征泛函

## §1　随机变量的特征函数, 反演公式

**定义 1.1**　设 $F(x)$ 是定义在 $\mathbf{R}$ 上的非负的、单调非降的右连续的有界函数, 即第一章 §4 中所定义的非负的测度有界的 (L-S) 函数. $\mu_F$ 是 $F(x)$ 所产生的 (L-S) 测度. (在无混淆的情况下, 本章简记 $\mu_F$ 为 $F$.) 称

$$f(t)=\int_{\mathbf{R}} e^{itx}F(dx)\quad (t\in\mathbf{R}) \tag{1.1}$$

为 $F(x)$ (或其对应的 (L-S) 测度 $F$) 的 Fourier 变换. 特别地, 称分布函数 (即标准的 (L-S) 函数) 的 Fourier 变换为**特征函数**, 或该分布函数所对应的随机变量的特征函数. 特征函数用 c.f. 表示, 分布函数用 d.f. 表示, 随机变量用 R.V. 表示.

显然, 非负的测度有界的 (L-S) 函数唯一地决定了其 Fourier 变换. 现在问: 逆命题是否成立? 回答是肯定的, 这就是下面的

**定理 1.1 (反演公式)**　设 $F(x)$ 是非负的测度有界的 (L-S) 函数, $f(t)$ 和 $F$ 分别是其 Fourier 变换及相应的 (L-S) 测度, 则对任意实数 $a<b$, 恒有:

$$\begin{aligned}&\frac{1}{2}[F(b)+F(b-0)]-\frac{1}{2}[F(a)+F(a-0)]\\&=\frac{1}{2\pi}\lim_{T\to\infty}\int_{-T}^{T}\frac{e^{-itb}-e^{-ita}}{-it}f(t)dt.\end{aligned} \tag{1.2}$$

**证** 令

$$I_T = \frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-itb}-e^{-ita}}{-it}f(t)dt,$$

则由

$$\begin{aligned}&\left|\frac{e^{-itb}-e^{-ita}}{-it}\bullet e^{itx}\right|\\ \leqslant &\left|\frac{(x-b)t-(x-a)t}{t}\right| = b-a\end{aligned}$$

可知下述重积分可交换次序:

$$\begin{aligned}I_T &= \frac{1}{2\pi}\int_{-T}^{T}\left(\int_{\mathbf{R}}\frac{e^{-itb}-e^{-ita}}{-it}e^{itx}F(dx)\right)dt\\ &= \frac{1}{2\pi}\int_{\mathbf{R}}\left(\int_{-T}^{T}\frac{e^{-it(b-x)}-e^{-it(a-x)}}{-it}dt\right)F(dx)\end{aligned}$$

令

$$J(u,T,x) = \int_{-T}^{T}\frac{e^{-it(u-x)}}{-it}dt,$$

则

$$\begin{aligned}I_T &= \frac{1}{2\pi}\int_{\mathbf{R}}(J(b,T,x)-J(a,T,x))F(dx)\\ &= \frac{1}{\pi}\int_{\mathbf{R}}\left(\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du\right)F(dx).\end{aligned}$$

再令

$$H(x) = \lim_{T\to\infty}\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du,$$

则由

$$\int_0^{\infty}\frac{\sin u}{u}du = \frac{\pi}{2}$$

可知

$$H(x) = \begin{cases}\pi, & \text{当 } a<x<b,\\ \dfrac{\pi}{2}, & \text{当 } x=a \text{ 或 } x=b,\\ 0, & \text{当 } x>b \text{ 或 } x<a,\end{cases}$$

所以

$$
\begin{aligned}
&\frac{1}{\pi}\int_{\mathbf{R}} H(x)F(dx) \\
&= \frac{1}{\pi}\left[\int_{(a,b)} \pi F(dx) + \int_{\{a\}} \frac{\pi}{2}F(dx) + \int_{\{b\}} \frac{\pi}{2}F(dx)\right] \\
&= \frac{1}{\pi}\left[\pi(F(b-0)-F(a)) + \frac{\pi}{2}(F(a)-F(a-0)) + \frac{\pi}{2}(F(b)-F(b-0))\right] \\
&= \frac{F(b)+F(b-0)}{2} - \frac{F(a)+F(a-0)}{2}.
\end{aligned}
$$

若能证

$$
\begin{aligned}
\lim_{T\to\infty} I_T &= \lim_{T\to\infty}\frac{1}{\pi}\int_{\mathbf{R}}\left(\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du\right)F(dx) \\
&= \frac{1}{\pi}\int_{\mathbf{R}}\left[\lim_{T\to\infty}\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du\right]F(dx),
\end{aligned}
$$

则定理 1.1 证毕. 而上式能积分号下取极限的充分条件是

$$
\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du
$$

是 $x$ 的有界函数. 但是根据第二中值定理可知:

$$
\left|\int_1^t \frac{\sin u}{u}du\right| \leqslant 2, \quad \left|\int_0^1 \frac{\sin u}{u}du\right| \leqslant 3,
$$

又因 $\dfrac{\sin u}{u}$ 是偶函数, 所以

$$
\int_{T(x-b)}^{T(x-a)}\frac{\sin u}{u}du
$$

是 $x$ 的有界函数. 定理证毕.

**推论 1.1** 在定理 1.1 的条件下, 若 $a, b$ 都是 $F(x)$ 的连续点, 即 $(a, b]$ 是 $F$ 的连续区间, 则有

$$
\begin{aligned}
F((a,b]) &= F(b) - F(a) \\
&= \lim_{T\to\infty} I_T.
\end{aligned} \tag{1.3}
$$

而 (L-S) 测度 $F$ 由其连续区间上的测度值所唯一决定, 又因为在定理 1.1 的条件下, $F(b) = \lim\limits_{a\to-\infty} F((a,b])$, 所以 $F(x)$ 和 $F$ 均由其 Fourier 变换 (特征函数) $f(t)$ 所唯一决定.

易见:

(1) 若 R.V.$X$ 是连续型的, 即有密度函数 $p(x)$, 则 $X$ 的 c.f. 为 $f(t) = \int_{\mathbf{R}} e^{itx}p(x)dx$;

(2) 若 R. V.$X$ 是离散型的, 即有分布 $\{p_k = P(X = x_k), k \geqslant 0\}$, 则 $X$ 的 c.f. 为 $f(t) = \sum\limits_{k\geqslant 0} e^{itx_k}p_k$;

(3) $X$ 的 c.f.$f(t) = \boldsymbol{E}(e^{itX})$ 恒存在.

**例 1.1** 若 d.f. $F(x)$ 对应的 (L-S) 测度 $F$ 只在一点 $a$ 上有正测度 $F(\{a\}) = 1$, 则称 $F(x)$ 为退化 d.f., 记之以 $\varepsilon_a(x)$. 称 $\varepsilon_0(x)$ 为零—律. $\varepsilon_a(x)$ 的特征函数为 $e^{ita}$. 退化分布的特征函数为常数 1.

**例 1.2** 若 R.V. $X$ 服从 **Poisson 分布**

$$P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots,$$

则 $X$ 的 c.f. 为

$$f(t) = e^{\lambda(e^{it}-1)}.$$

**例 1.3** 若 R.V. $X$ 服从**二项分布**:

$$P(X = k) = \binom{n}{k}p^k(1-p)^{n-k}, \quad k = 0, 1, \cdots, n,$$

则 $X$ 的 c.f. 为

$$f(t) = (pe^{it} + (1-p))^n.$$

**例 1.4** 若 R.V. $X$ 服从**标准正态分布** (记之为 $N(0,1)$), 即 $X$ 有密度函数

$$p(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2},$$

则 $X$ 的 c.f. $f(t) = e^{-\frac{1}{2}t^2}$.

事实上, 由于

$$f(t) = \int_{\mathbf{R}} \frac{1}{2\pi} \cdot e^{itx-\frac{1}{2}x^2}dx, \tag{1.4}$$

而且

$$\left|\frac{\partial}{\partial x}\left(\frac{1}{\sqrt{2\pi}}e^{itx-\frac{1}{2}x^2}\right)\right| = \left|\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2} \cdot it\right|,$$

$$\int_{\mathbf{R}}\left|\frac{1}{\sqrt{2\pi}}xe^{-\frac{1}{2}x^2}\right|dx < \infty,$$

所以对 $f(t)$ 的积分表达式 (1.4), 可以在积分号下取导数, 得:

$$
\begin{aligned}
f'(t) &= \int_{\mathbf{R}} -\frac{1}{\sqrt{2\pi}} \cdot x e^{-\frac{1}{2}x^2} \sin tx dx \\
&= -t \int_{\mathbf{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cos tx dx \\
&= -t f(t).
\end{aligned}
$$

所以

$$\log f(t) = -\frac{1}{2}t^2 + C.$$

由 $f(0) = 1$ 得 $C = 0$, 从而 $f(t) = e^{-\frac{1}{2}t^2}$.

更一般地, 若 $X$ 服从 $N(a, \sigma^2)$, 即期望为 $a$, 方差为 $\sigma^2$ 的正态分布, 则其 c.f. 为

$$f(t) = e^{iat - \frac{1}{2}\sigma^2 t^2}.$$

**例 1.5** 若存在实数 $a$ 及正数 $u$, 使 $P(X = a + ku) = p_k$, $k = 0, \pm1, \pm2, \cdots$, $p_k \geqslant 0, \sum\limits_k p_k = 1$, 则称 R.V.$X$ 服从**格子点分布**. 令 $f(t)$ 是 $X$ 的 c.f., 称 $u$ 为其间隔.

R.V. $X$ 服从格子点分布的充分必要条件是存在一个 $t_0 \neq 0$, 使 $|f(t_0)| = 1$.

**证** 必要性显然成立.

充分性. 若存在 $t_0 \neq 0$, 使 $|f(t_0)| = 1$, 则存在实数 $a$, 使 $f(t_0) = e^{iat_0}$. 所以

$$1 = e^{-iat_0} f(t_0) = \boldsymbol{E}(e^{-iat_0 + iXt_0}),$$

从而

$$\boldsymbol{E}(\cos t_0(X - a)) = 1,$$

由 $|\cos t_0(X - a)| \leqslant 1$ 得

$$P(1 - \cos t_0(X - a) = 0) = 1,$$

亦即

$$P\left(\bigcup_{k=0,\pm1,\cdots} \{t_0(X - a) = 2k\pi\}\right) = 1.$$

令

$$p_k = P\left(X = a + \frac{2k\pi}{|t_0|}\right), \quad k = 0, \pm1, \cdots,$$

则 $X$ 服从格子点分布 $\{p_k, k = 0, \pm1, \pm2, \cdots\}$, 其间隔 $u = 2\pi/|t_0|$.

**推论 1.2** 若存在 $t_0 \neq 0, t_1 \neq 0, t_0/t_1$ 不是有理数使 $|f(t_0)| = |f(t_1)| = 1$, 则 $f(t)$ 对应的 R.V.$X$ 服从退化分布.

**证** 设 $|f(t_0)| = |f(t_1)| = 1$, 则由例 1.5, 若令

$$p_k = P\left(X = a_0 + \frac{2k\pi}{|t_0|}\right), \quad k = 0, \pm 1, \pm 2, \cdots,$$
$$q_k = P\left(X = a_1 + \frac{2k\pi}{|t_1|}\right), \quad k = 0, \pm 1, \pm 2, \cdots,$$

则 $p_k \geqslant 0, q_k \geqslant 0, \sum\limits_k p_k = \sum\limits_k q_k = 1$. 谬设 $X$ 不服从退化分布, 则 $p_{k_1} > 0, p_{k_2} > 0, p_{k_1} = q_{n_1}, p_{k_2} = q_{n_2}$. 所以

$$\left|(k_1 - k_2)\frac{2\pi}{t_0}\right| = \left|(n_1 - n_2)\frac{2\pi}{t_1}\right|,$$

即 $t_0$ 与 $t_1$ 之比为有理数. 这与假设矛盾, 所以 $X$ 服从退化分布.

**定理 1.2** 随机变量的特征函数有下述性质:

(1) R.V.$X$ 的 c.f.$f(t)$ 在 $\mathbf{R}$ 上一致连续, $|f(t)| \leqslant f(0) = 1, f(-t) = \overline{f(t)}$ (此处 $\overline{f}$ 是 $f$ 的复共轭);

(2) 若 R.V.$X$ 的特征函数为 $f(t)$, 则 $Y = a + bX$ 的 c.f. 为 $g(t) = e^{iat}f(bt)$, 此处 $a, b$ 为任意二实数.

(3) 若 d.f.$F_1(x), F_2(x)$ 所对应的 c.f. 分别为 $f_1(t), f_2(t)$, 则 $F_1(x)$ 和 $F_2(x)$ 所对应的 (L-S) 测度的卷积 $F_1 * F_2$ 的特征函数为 $f_1(t)f_2(t)$. 反之也对.

(4) 随机变量 $X$ 服从对称分布 (即 $P(X \leqslant x) = P(-X \leqslant x)$) 的充分必要条件是: $X$ 的特征函数是实值的.

证明甚易, 读者可作为习题验证之.

## §2 连续性定理

在第 1 节中, 我们看到了分布函数与其特征函数的一一对应关系. 在这一节中, 我们要研究分布函数列 $\{F_n(x), n \geqslant 1\}$ 的弱收敛与其对应的特征函数列 $\{f_n(t), n \geqslant 1\}$ 的某种收敛性的关系.

**定理 2.1** 设分布函数列 $\{F(x), F_n(x), n \geqslant 1\}$ 所对应的特征函数列为 $\{f(t), f_n(t), n \geqslant 1\}$. 若 $F_n(x) \xrightarrow{w} F(x)$, 则 $f_n(t) \to f(t)$ 在 $t$ 属于任何有限区间 $I$ 上一致成立.

**证** 这可由第三章定理 5.6 直接推出.

**定理 2.2** 设分布函数列 $\{F_n(x), n \geqslant 1\}$ 所对应的特征函数列为 $\{f_n(t), n \geqslant 1\}$. 若

$$\lim_{n\to\infty} f_n(t) = f(t), \quad \text{L}-\text{a.s.}, \quad (\text{L 表示 Lebesgue 测度}),$$

则存在唯一一个 $\mathscr{B}(\mathbf{R})$ 上的有限测度 $F$, 使

(1) $F_n \xrightarrow{\text{lw}} F$, 其中 $F_n$ 是 $F_n(x)$ 所产生的 (L-S) 测度;

(2) $f(t) = \int_{\mathbf{R}} e^{itx} F(dx)$.

**证**　由第三章推论 5.1, 为证 (1), 只需证明 $\{F_n, n \geqslant 1\}$ 的任何一个局部弱收敛子列都收敛到同一个极限.

任取 $\{F_n\}$ 的一个局部弱收敛子列 $\{G_n\}$,

$$G_n \overset{\text{lw}}{\Longrightarrow} F. \tag{2.1}$$

往证 $F$ 与 $\{G_n\}$ 之选取无关. 事实上, 由 (2.1) 及 $G_n(\mathbf{R}) \equiv 1$ 得知 $F$ 在其任一连续区间上的测度值 $\leqslant 1$, 又因为

$$\lim_{|x|\to\infty} \frac{e^{itx}-1}{ix} = 0 \quad \text{对 } t \in \mathbf{R} \text{ 一致成立},$$

所以由第三章定理 5.10 得

$$\lim_{n\to\infty} \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G_n(dx) = \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} F(dx). \tag{2.2}$$

令 $g_n$ 为 $G_n$ 对应之 c.f., 则 $\{g_n, n \geqslant 1\}$ 是 $\{f_n, n \geqslant 1\}$ 的子序列, 所以

$$\lim_{n\to\infty} g_n(t) = f(t), \quad \text{L}-\text{a.s.}. \tag{2.3}$$

因此

$$\lim_{n\to\infty} \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G_n(dx) = \lim_{n\to\infty} \int_0^t g_n(u)du = \int_0^t f(u)du.$$

以上式代入 (2.2) 得

$$\int_0^t f(u)du = \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} F(dx). \tag{2.4}$$

对 (2.4) 求导数得:

$$f(t) = \int_{\mathbf{R}} e^{itx} F(dx), \quad \text{L}-\text{a.s.}.$$

用定理 1.1, $F$ 由 $f$ 唯一决定而与子序列 $\{G_n\}$ 的选取无关. 这就证明了 (1). 而在上述证明中, 还说明了 $F$ 的 Fourier 变换与 $f$ 是 L−a.s. 相等的. 定理证毕.

**定理 2.3** 设 d.f. $F_n(x)$ 的 c.f. 为 $f_n(t),(n \geqslant 1)$, 若

$$\lim_{n\to\infty} f_n(t) = f(t), \quad 且\ f(t)\ 在\ 0\ 连续,$$

则存在 d.f. $F(x)$ 使

(1) $F_n(x) \xrightarrow{w} F(x)$;

(2) $f(t) = \int_{\mathbf{R}} e^{itx} F(dx)$;

(3) $\lim\limits_{n\to\infty} f_n(t) = f(t)$, 在 $t$ 属于任何有界区间上一致成立.

**证** 由 $\lim\limits_{n\to\infty} f_n(t) = f(t)$ 及定理 2.2 得知

$$F_n \overset{\text{lw}}{\Longrightarrow} F \tag{2.5}$$

且

$$f(t) = \int_{\mathbf{R}} e^{itx} F(dx), \quad \text{L}-\text{a.s.}.$$

但 $f(t)$ 在 0 连续, 所以

$$f(0) = \int_{\mathbf{R}} F(dx). \tag{2.6}$$

而

$$f(0) = \lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} F_n(\mathbf{R}), \tag{2.7}$$

由 (2.6) 和 (2.7) 得知: 任给 $\varepsilon > 0$, 存在正整数 $n_0$ 及区间 $I_0$, 使得

$$F_n(\mathbf{R}) - F_n(I) < \varepsilon \quad (\forall n \geqslant n_0, I \supset I_0). \tag{2.8}$$

由 (2.5)、(2.8) 及第三章定理 5.7 知

$$F_n \overset{\text{w}}{\Longrightarrow} F.$$

这就证明了 (1). 由 (1) 及定理 2.1, 定理 2.3 的后二结论成立. 定理证毕.

**定理 2.4** 设 d.f. $F_n(x)$ 对应之 c.f. 为 $f_n(t)(n \geqslant 1)$. 若 $\lim\limits_{n\to\infty} f_n(t) = f(t)$ (当 $|t| < \delta$), 且 $f(t)$ 在 $t = 0$ 连续, 则

(1) $\{F_n\}$ 的任何一个局部弱收敛子列 $\{G_n\}$ 都弱收敛;

(2) $\{F_n\}$ 的弱收敛子列的极限所对应的特征函数在 $|t| < \delta$ 上与 $f(t)$ 相等;

(3) $\lim\limits_{n\to\infty} f_n(t) = f(t)$ 在 $|t| < \delta$ 上一致成立.

**证** (1) 令 $G_n \overset{\text{lw}}{\Longrightarrow} G, G_n$ 对应之 Fourier 变换为 $g_n(t)$, 则由第三章定理 5.10 知

$$\lim_{n\to\infty} \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G_n(dx) = \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G(dx). \tag{2.9}$$

但是

$$\lim_{n\to\infty} f_n(t) = f(t), \quad |t| < \delta,$$

$\{g_n\}$ 是 $\{f_n\}$ 的子序列, 所以

$$\lim_{n\to\infty}\int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G_n(dx) = \lim_{n\to\infty}\int_0^t g_n(u)du$$
$$= \int_0^t f(u)du, \quad |t| < \delta.$$

把上式代入 (2.9) 得

$$\int_0^t f(u)du = \int_{\mathbf{R}} \frac{e^{itx}-1}{ix} G(dx), \quad |t| < \delta. \tag{2.10}$$

对 (2.10) 求导数得:

$$f(t) = \int_{\mathbf{R}} e^{itx} G(dx), \quad |t| < \delta.$$

又因为 $f(t)$ 在 $t=0$ 连续, 故

$$f(0) = G(\mathbf{R}), \tag{2.11}$$

$$f(0) = \lim_{n\to\infty} g_n(0) = \lim_{n\to\infty} G_n(\mathbf{R}), \tag{2.12}$$

比较 (2.11) 和 (2.12) 并注意 $G_n \stackrel{\text{lw}}{\Longrightarrow} G$, 再应用第三章定理 5.3 可知

$$G_n \stackrel{\text{w}}{\Longrightarrow} G.$$

(2) 若 $\{G_n\}$ 是 $\{F_n\}$ 的弱收敛子列. $G_n \stackrel{\text{w}}{\Longrightarrow} G$. (这样的 $\{G_n\}$ 由第三章定理 5.6 以及已证明的 (1) 得知必存在.) 由第三章定理 5.9 知:

$$\int_{\mathbf{R}} e^{itx} G(dx) = \lim_{n\to\infty}\int_{\mathbf{R}} e^{itx} G_n(dx)$$
$$= \lim_{n\to\infty} g_n(t) = f(t) \quad (\text{在 } |t| < \delta \text{ 一致成立}), \tag{2.13}$$

此即 (2) 成立.

(3) 反设 (3) 不成立, 则存在 $\varepsilon > 0$, 及正整数列 $\{n_k, k \geqslant 1\}$ 使

$$\sup_{|t|<\delta} |f_{n_k}(t) - f(t)| > \varepsilon \quad (\forall k \geqslant 1). \tag{2.14}$$

由于 $\{f_{n_k}, k \geqslant 1\}$ 是 $\{f_n, n \geqslant 1\}$ 的子序列, 所以由本定理的 (1) 和 (2), $\{F_{n_k}, k \geqslant 1\}$ 又有子序列 $\{G_{n_k}, k \geqslant 1\}$ 使

$$G_{n_k} \stackrel{\text{w}}{\Longrightarrow} G^*,$$

$$\lim_{k\to\infty} g_{n_k}(t) = f(t) \ \ (\text{在 } |t| < \delta \text{ 上一致成立}), \tag{2.15}$$

其中 $g_{n_k}$ 是 $G_{n_k}$ 所对应的 Fourier 变换.

而 (2.14) 与 (2.15) 是矛盾的, 所以 (3) 成立. 定理证毕.

**推论 2.1** 若 $f(t), f_n(t)$ 都是 c.f.$(n \geqslant 1)$, 而且 $f_n(t) \to f(t)$, 则 $f_n(t) \to f(t)$ 在 $t$ 的任何有限区间上一致成立.

**证** 由定理 2.4 立即可得.

**推论 2.2** 若 $f(t), f_n(t)$ 都是 c.f.$(n \geqslant 1), f_n(t) \to f(t), t_n \to t_0$, 则 $f_n(t_n) \to f(t_0)$.

**证** $|f_n(t_n) - f(t_0)| \leqslant |f_n(t_n) - f(t_n)| + |f(t_n) - f(t_0)|$,

所以, 由 $f(t)$ 的连续性及推论 2.1 即得推论 2.2.

**推论 2.3** 设分布函数列 $\{F(x), F_n(x), n \geqslant 1\}$ 对应的特征函数列为 $\{f(t), f_n(t), n \geqslant 1\}$, 则 $F_n(x) \xrightarrow{\text{w}} F(x)$ 的充分必要条件是: $f_n(t) \to f(t)$.

**证** 这是定理 2.1 和定理 2.3 的直接推论.

**定理 2.5** 设 R.V.$X$ 的 d.f. 为 $F(x)$, c.f. 为 $f(t), \alpha$ 是 $X$ 的中位数, 即 $F(\alpha - 0) \leqslant \dfrac{1}{2}, F(\alpha) \geqslant \dfrac{1}{2}$, 则恒有下列不等式:

(1) 增量不等式

$$|f(t) - f(t+h)|^2 \leqslant 2f(0)[f(0) - \mathscr{R}(f(h))], \tag{2.16}$$

$$1 - \mathscr{R}(f(2t)) \leqslant 4[1 - \mathscr{R}(f(t))], \tag{2.17}$$

其中 $\mathscr{R}(f)$ 表 $f$ 的实部.

(2) 积分不等式

设 $\alpha$ 是中位数, $F_\alpha(x) = F(x + \alpha), A \subset [0, \delta]$, 且 $A$ 的 Lebesgue 测度 $L(A) = \rho > 0$, 则

$$\int_A (1 - |f(t)|^2) dt \geqslant \frac{\rho^3}{4(\delta^2 + 4\pi^2)} \int_{\mathbf{R}} \frac{x^2}{1+x^2} F_\alpha(dx). \tag{2.18}$$

特别地, 若 $A = [0, t]$, 则 (2.17) 还可加强为

$$\begin{aligned} m(t) \int_0^t [1 - \mathscr{R}(f(u))] du &\leqslant \int_{\mathbf{R}} \frac{x^2}{1+x^2} F(dx) \\ &\leqslant M(t) \int_0^t [1 - \mathscr{R}(f(u))] du, \end{aligned} \tag{2.19}$$

其中

$$\frac{1}{m(t)}=\sup_x\left[t\left(1-\frac{\sin tx}{tx}\right)\frac{1+x^2}{x^2}\right],$$
$$\frac{1}{M(t)}=\inf_x\left[t\left(1-\frac{\sin tx}{tx}\right)\frac{1+x^2}{x^2}\right].$$

(3) 截尾不等式

众所周知: 若 R.V.$X$ 的 c.f. 为 $f(t)$, 而 $X$ 的数学期望 $m\stackrel{\text{def}}{=\!=}\boldsymbol{E}(X)$ 存在时, 在概率极限理论中, 常利用 "搬中心" 技巧, 即用 $X-m$ 来代替 $X$, 而 $X-m$ 的特征函数为 $e^{-imt}f(t)$. 但当 $\boldsymbol{E}(X)$ 不存在时, 人们就考虑用

$$a(\tau)\stackrel{\text{def}}{=\!=}\int_{|x|<\tau}xF(dx) \tag{2.20}$$

来代替 $\boldsymbol{E}(X)$. (因为 $a(\tau)$ 总是存在的.) 于是, 考虑 $|e^{-ia(\tau)t}f(t)-1|$ 的各种近似估计就有必要. 我们恒有

$$\begin{aligned}&\left|e^{-ia(\tau)t}f(t)-1\right|\\&\leqslant(2+|t|\tau)\int_{|x|\geqslant\tau}F(dx)+\frac{t^2}{2}\int_{|x|<\tau}(x-a(\tau))^2F(dx).\end{aligned} \tag{2.21}$$

$$\begin{aligned}&\left|e^{-ia(\tau)t}f(t)-1\right|\\&\leqslant\left[(|\alpha|+\tau)\tau t^2+\tau|t|+2\right]\int_{|x|\geqslant\tau}F(dx)\\&\quad+\frac{t^2}{2}\int_{|x|<\tau}(x-\alpha)^2F(dx),\end{aligned} \tag{2.22}$$

其中 $\alpha$ 是 $X$ (或者说是 $F(x)$) 的中位数.

$$\begin{aligned}&\left|e^{-ia(\tau)t}f(t)-1\right|\\&\leqslant A(t,\tau)\int_{\mathbf{R}}\frac{(x-a(\tau))^2}{1+(x-a(\tau))^2}F(dx).\end{aligned} \tag{2.23}$$

当 $|\alpha|\leqslant\dfrac{\tau}{2}$ 时, 还有

$$\begin{aligned}&\int_{\mathbf{R}}\frac{(x-a(\tau))^2}{1+(x-a(\tau))^2}F(dx)\\&\leqslant B(\tau,\delta)\int_0^\delta(1-|f(t)|^2)dt\quad\left(\text{当 }|\alpha|<\frac{\tau}{2}\right).\end{aligned} \tag{2.24}$$

若用 $F_\alpha(x)\stackrel{\text{def}}{=\!=}F(x+\alpha), f_\alpha(t)\stackrel{\text{def}}{=\!=}e^{-i\alpha t}f(t)$ 来替代 $F(x)$ 和 $f(t)$, 我们还有

下列相应的不等式:

$$\left|e^{-ib(\tau)t}f(t)-1\right| \leqslant (1+\tau^2)\left(\frac{3|t|^2}{2}+\frac{|t|}{\tau}+\frac{2}{\tau^2}\right)\frac{4(\delta^2+4\pi^2)}{\rho^3}\cdot \int_A (1-|f(t)|^2)dt, \tag{2.25}$$

其中

$$b(\tau)=\alpha+a_\alpha(\tau),$$
$$a_\alpha(\tau)=\int_{|x|<\tau} xF_\alpha(dx).$$

还有不等式:

$$\int_{\mathbf{R}} \frac{(x-b(\tau))^2}{1+(x-b(\tau))^2}F(dx) \leqslant 2(1+4\pi^2)(1+\tau^2)\left(\frac{1}{2}+\frac{1}{2\tau}+\frac{2}{\tau^2}\right)\frac{4(\delta^2+4\pi^2)}{\rho^3}\cdot\int_A (1-|f(t)|^2)dt. \tag{2.26}$$

(4) 几个积分的估计

$$\int_{|x|\geqslant\tau} F(dx) \leqslant \frac{2(\delta^2+4\pi^2)}{\rho^3}\cdot\frac{1+\tau^2}{\tau^2}\int_A \mathscr{R}(1-f(t))dt. \tag{2.27}$$

$$\int_{|x-\alpha|\geqslant\tau} F(dx) \leqslant \frac{4(\delta^2+4\pi^2)}{\rho^3}\cdot\frac{1+\tau^2}{\tau^2}\int_A (1-|f(t)|^2)dt. \tag{2.28}$$

$$\int_{|x-b(\tau)|\geqslant\tau_1} F(dx) \leqslant \frac{1+\tau_1^2}{\tau_1^2}\cdot 2\cdot(1+4\pi^2)(1+\tau^2)\cdot \left(\frac{1}{2}+\frac{1}{2\tau}+\frac{1}{\tau^2}\right)\cdot\frac{4(\delta^2+4\pi^2)}{\rho^3}\int_A (1-|f(t)|^2)dt. \tag{2.29}$$

$$\int_{|x|<\tau} x^2F(dx) \leqslant (1+\tau^2)\frac{2(\delta^2+4\pi^2)}{\rho^3}\int_A \mathscr{R}(1-f(t))dt. \tag{2.30}$$

$$\int_{|x-\alpha|<\tau}(x-\alpha)^2F(dx)-\left(\int_{|x-\alpha|<\tau}(x-\alpha)F(dx)\right)^2 \leqslant k(\tau,\rho,\delta)\int_A (1-|f(t)|^2)dt, \tag{2.31}$$

其中 $k(\tau,\rho,\delta)$ 是依赖 $\tau,\rho,\delta$ 的常数.

此定理中的诸不等式都是在利用特征函数作工具来证明一些概率极限理论时才用. 其证明比较繁琐, 故在此略去, 有兴趣的读者可参阅文献 [36] P.43–53.

## §3 特征函数的 Taylor 展式

在用特征函数作工具研究概率极限理论时, 经常要用特征函数的 Taylor 展式作估计, 这就需要知道特征函数具有多少阶导数, 而特征函数所对应的随机变量有几阶矩存在, 与特征函数有几阶导数是有密切关系的.

**定理 3.1** 设 $f(t)$ 是 R.V.$X$ 的 c.f.. 若 $f(t)$ 在 0 点的 $2n$ 阶导数 $f^{(2n)}(0)$ 存在, 则对任何 $0 \leqslant r \leqslant 2n, \boldsymbol{E}(|X|^r) < \infty$.

**证** 令 $\Delta_h(f(0)) = f(h) - f(-h)$

$$\begin{aligned}\Delta_h^{(k)}(f(0)) &= \Delta_h\left(\Delta_h^{(k-1)}(f(0))\right),\\ &= \sum_{j=0}^{k}(-1)^j\binom{k}{j}f((k-2j)h), \quad k \geqslant 2.\end{aligned}$$

再令 $F(x)$ 为 $X$ 的 d.f., 则

$$\begin{aligned}f^{(2n)}(0) &= \lim_{h\to 0}\frac{\Delta_h^{(2n)}(f(0))}{(2h)^{2n}}\\ &= \lim_{h\to 0}\frac{\sum\limits_{j=0}^{2n}(-1)^j\dbinom{2n}{j}f((2n-2j)h)}{(2h)^{2n}}\\ &= \lim_{h\to 0}\frac{\sum\limits_{j=0}^{2n}(-1)^j\dbinom{2n}{j}\displaystyle\int_{\mathbf{R}}e^{i(2n-2j)hx}F(dx)}{(2h)^{2n}}\\ &= \lim_{h\to 0}\int_{\mathbf{R}}\left(\frac{e^{ihx}-e^{-ihx}}{2h}\right)^{2n}F(dx)\\ &= \lim_{h\to 0}\int_{\mathbf{R}}\left(\frac{2i\sin hx}{2h}\right)^{2n}F(dx)\\ &= \lim_{h\to 0}\int_{\mathbf{R}}(-1)^n x^{2n}\left(\frac{\sin hx}{hx}\right)F(dx).\end{aligned}$$

但是, 由 Fatou 引理知

$$\begin{aligned}\boldsymbol{E}(|X|^{2n}) &= \int_{\mathbf{R}}\lim_{h\to 0}x^{2n}\left(\frac{\sin hx}{hx}\right)^{2n}F(dx)\\ &\leqslant \lim_{h\to 0}\int_{\mathbf{R}}x^{2n}\left(\frac{\sin hx}{hx}\right)^{2n}F(dx).\end{aligned}$$

若 $f^{(2n)}(0)$ 存在, 则

$$\lim_{h\to 0}\int_{\mathbf{R}}x^{2n}\left(\frac{\sin hx}{hx}\right)^{2n}F(dx) = |f^{(2n)}(0)| < \infty,$$

因此 $\boldsymbol{E}(|X|^{2n}) < \infty$. 再利用 Hölder 不等式知 $\boldsymbol{E}(|X|^r) < \infty$ ($\forall 0 \leqslant r \leqslant 2n$). 定理证毕.

**定理 3.2** 设 R.V. $X$ 的 c.f. 为 $f(t)$. 若 $\boldsymbol{E}(|X|^n) < \infty$, 则 $f(t)$ 有 $n$ 阶连续导数, 且

$$f^{(n)}(t) = \int_{\mathbf{R}} (ix)^n e^{itx} F(dx),$$

其中 $n$ 为任意正整数, $F(x)$ 为 $X$ 的 d.f..

**证** 对 $n$ 作归纳法. 若 $\boldsymbol{E}(|X|) < \infty$, 则用 Lebesgue 控制收敛定理可证:

$$f'(t) = \int_{\mathbf{R}} \frac{\partial}{\partial t}(e^{itx}) F(dx) = \int_{\mathbf{R}} ixe^{itx} F(dx)$$

是 $t$ 的连续函数.

设定理对 $n = k$ 成立. 要证它对 $n = k+1$ 也成立. 事实上, 若定理对 $n = k$ 成立, 即是当 $\boldsymbol{E}(|X|^k) < \infty$ 时有

$$f^{(k)}(t) = \int_{\mathbf{R}} (ix)^k e^{itx} F(dx)$$

是 $t$ 的连续函数. 今设 $\boldsymbol{E}(|X|^{k+1}) < \infty$. 再用 Lebesgue 控制收敛定理可证

$$f^{(k+1)}(t) = \int_{\mathbf{R}} (ix)^{k+1} e^{itx} F(dx)$$

是 $t$ 的连续函数. 定理证毕.

下面我们讨论特征函数的 Taylor 展开. 记 $m_n = \boldsymbol{E}(|X|^n)$.

**定理 3.3** 若 $m_n < \infty$, 则特征函数 $f(t)$ 有:

$$\begin{aligned} f(t) &= \sum_{k=0}^{n-1} f^{(k)}(0) \frac{t^k}{k!} + O(t^n). \\ &= \sum_{k=0}^{n} f^{(k)}(0) \frac{t^k}{k!} + o(t^n) \quad (t \to 0). \end{aligned}$$

**证** 由于 $m_n < \infty$, 所以 $f(t)$ 有 $n$ 阶连续导数, 故由 Taylor 展开公式有

$$\begin{aligned} f(t) &= \sum_{k=0}^{n-1} f^{(k)}(0) \frac{t^k}{k!} + \int_0^t f^{(n)}(s) \frac{(t-s)^{n-1}}{(n-1)!} ds \\ &= \sum_{k=0}^{n} f^{(k)}(0) \frac{t^k}{k!} + \int_0^t (f^{(n)}(s) - f^{(n)}(0)) \frac{(t-s)^{n-1}}{(n-1)!} ds. \end{aligned}$$

令

$$R_{n-1} = \int_0^t f^{(n)}(s)\frac{(t-s)^{n-1}}{(n-1)!}ds,$$
$$\widetilde{R}_n = \int_0^t (f^{(n)}(s) - f^{(n)}(0))\frac{(t-s)^{n-1}}{(n-1)!}ds,$$

则由 $f^{(n)}(t) = \int_{\mathbf{R}}(ix)^n e^{itx}F(dx)$ 得

$$|f^{(n)}(t)| \leqslant \int_{\mathbf{R}} |x|^n F(dx) = m_n < \infty,$$

所以

$$|R_{n-1}| \leqslant m_n \left|\int_0^t \frac{(t-s)^{n-1}}{(n-1)!}ds\right| = m_n \frac{|t|^n}{n!} = O(t^n) \quad (t\to 0),$$
$$|\widetilde{R}_n| \leqslant \sup_{0\leqslant s\leqslant t} |f^{(n)}(s) - f^{(n)}(0)| \left|\int_0^t \frac{(t-s)^{n-1}}{(n-1)!}ds\right|$$
$$\leqslant \sup_{0\leqslant s\leqslant t} |f^{(n)}(s) - f^{(n)}(0)|\frac{|t|^n}{n!} = o(t^n), \quad (t\to 0).$$

定理得证.

**定理 3.4** 若 $m_{n+\delta} < \infty$ $(0<\delta<1)$, 则

$$f(t) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!}t^k + \theta 2^{1-\delta} m_{n+\delta} \frac{t^{n+\delta}}{(1+\delta)(2+\delta)\cdots(n+\delta)},$$

其中 $|\theta| \leqslant 1$.

**证** 由定理 3.3 可知:

$$f(t) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!}t^k + \widetilde{R}_n,$$
$$|\widetilde{R}_n| = \left|\int_0^t |f^{(n)}(s) - f^{(n)}(0)|\frac{(t-s)^{n-1}}{(n-1)!}ds\right.$$
$$= \left|\int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\left(\int_{\mathbf{R}}(ix)^n(e^{itx}-1)F(dx)\right)ds\right|.$$

但是

$$|e^{isx} - 1| \leqslant 2^{1-\delta}|sx|^\delta \quad (0<\delta<1),$$

所以

$$
\begin{aligned}
|\widetilde{R}_n| &\leqslant \left|\int_0^t \frac{(t-s)^{n-1}}{(n-1)!}\left(\int_{\mathbf{R}} |x|^n 2^{1-\delta}|sx|^{\delta} F(dx)\right) ds\right| \\
&\leqslant 2^{1-\delta} m_{n+\delta}\left|\int_0^t \frac{s^{\delta}(t-s)^{n-1}}{(n-1)!} ds\right| \\
&= 2^{1-\delta} m_{n+\delta}\left|\int_0^1 \frac{x^{\delta}(1-x)^{n-1}t^{n+\delta}}{(n-1)!} dx\right| \\
&\leqslant 2^{1-\delta} m_{n+\delta}\frac{|t|^{n+\delta}}{(n-1)!}\int_0^1 x^{\delta}(1-x)^{n-1} dx \\
&= 2^{1-\delta} m_{n+\delta}\frac{|t|^{n+\delta}}{(n-1)!}\mathrm{B}(\delta+1,n) \\
&= 2^{1-\delta} m_{n+\delta}\frac{|t|^{n+\delta}}{(n-1)!}\frac{\Gamma(\delta+1)\Gamma(n)}{\Gamma(\delta+1+n)} \\
&= 2^{1-\delta} m_{n+\delta}\frac{|t|^{n+\delta}}{(n-1)!}\cdot\frac{\Gamma(\delta+1)(n-1)!}{(n+\delta)(n-1+\delta)\cdots(1+\delta)\Gamma(1+\delta)} \\
&= 2^{1-\delta} m_{n+\delta}\frac{|t|^{n+\delta}}{(n+\delta)(n-1+\delta)\cdots(1+\delta)}.
\end{aligned}
$$

定理证毕.

## *§4 Khinchin-Bochner 定理

**定义 4.1** 称实变复值函数 $f(t)$ 是非负定的, 如果对任何正整数 $n$, 任何实数 $t_1, t_2, \cdots, t_n$ 及任何复数 $\xi_1, \cdots, \xi_n$, 都有

$$
\sum_{j=1}^{n}\sum_{k=1}^{n} f(t_j - t_k)\xi_j\overline{\xi}_k \geqslant 0, \tag{4.1}
$$

其中 $\overline{\xi}_k$ 是 $\xi_k$ 的复共轭.

易证非负定函数具有下述性质:

(1)
$$
f(0) \geqslant 0. \tag{4.2}
$$

事实上, 在 (4.1) 中取 $n=1, t_1=0, \xi_1=1$ 即可得 (4.2).

(2) 对任何实数 $t$, 有

$$
f(t) = \overline{f}(-t), \tag{4.3}
$$

其中 $\overline{f}$ 是 $f$ 的复共轭.

事实上, 在 (4.1) 中取 $n=2, t_1=0, t_2=1$ 得

$$\begin{aligned} 0 &\leqslant \sum_{j=1}^{2}\sum_{k=1}^{2} f(t_j-t_k)\xi_j\overline{\xi}_k \\ &= f(0)\left[|\xi_1|^2+|\xi_2|^2\right]+f(-t)\xi_1\overline{\xi}_2+f(t)\xi_2\overline{\xi}_1, \end{aligned} \tag{4.4}$$

所以 $f(-t)\xi_1\overline{\xi}_2+f(t)\xi_2\overline{\xi}_1$ 是实数. 令 $f(-t)=\alpha_1+i\beta_1, f(t)=\alpha_2+i\beta_2, \xi_1\overline{\xi}_2=\gamma+i\delta, \xi_2\overline{\xi}_1=\gamma-i\delta$. 再令 $\mathscr{I}(f)$ 为 $f$ 的虚部, 则

$$\begin{aligned} 0 &= \mathscr{I}(f(-t)\xi_1\overline{\xi}_2+f(t)\xi_2\overline{\xi}_1) \\ &= \alpha_1\delta+\beta_1\gamma-\alpha_2\delta+\beta_2\gamma, \end{aligned}$$

即

$$(\alpha_1-\alpha_2)\delta+(\beta_1+\beta_2)\gamma=0.$$

而 $\xi_1,\xi_2$ 可以任意, 即 $\gamma$ 和 $\delta$ 可以任意, 所以 $(\alpha_1-\alpha_2)=0, \beta_1+\beta_2=0$, 此即 $f(-t)-\overline{f}(t)=(\alpha_1-\alpha_2)+(\beta_1+\beta_2)\mathrm{i}=0$, 亦即 $\overline{f}(t)=f(-t)$.

(3) 对任意 $t$, 恒有 $|f(t)|\leqslant|f(0)|$.

事实上, 在 (4.4) 中令 $\xi_1=f(t), \xi_2=-|f(t)|$, 则

$$0\leqslant 2f(0)|f(t)|^2-2|f(t)||f(t)|^2. \tag{4.5}$$

若 $f(t)=0$, 则由 $f(0)\geqslant 0$ 得知 $f(0)\geqslant|f(t)|$; 若 $f(t)\neq 0$, 则将 (4.5) 两边除以 $|f(t)|^2$ 即得 $|f(t)|\leqslant f(0)$.

(4) 若非负定函数 $f(t)$ 满足 $f(0)=0$, 则 $f(t)\equiv 0$.

**定理 4.1 (Khinchin – Bochner)** 设 $f(t)$ 是连续函数, $f(0)=1$, 则 $f(t)$ 是非负定函数的充分必要条件是: $f(t)$ 是某个随机变量的特征函数.

**证** 充分性. 若存在 d.f.$F(x)$ 使

$$f(t)=\int_{\mathbf{R}} e^{itx}F(dx),$$

则对任意正整数 $n$, 任意实数 $t_1,\cdots,t_n$, 任意复数 $\xi_1,\cdots,\xi_n$, 都有

$$\begin{aligned} \sum_{j=1}^{n}\sum_{k=1}^{n} f(t_j-t_k)\xi_j\overline{\xi}_k &= \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\int_{\mathbf{R}} e^{i(t_j-t_k)x}F(dx)\right)\xi_j\overline{\xi}_k \\ &= \int_{\mathbf{R}}\left(\sum_{j=1}^{n}\sum_{k=1}^{n} e^{i(t_j-t_k)x}\xi_j\overline{\xi}_k\right)F(dx) \\ &= \int_{\mathbf{R}}\left|\sum_{j=1}^{n} e^{it_jx}\xi_j\right|^2 F(dx)\geqslant 0, \end{aligned}$$

所以 $f(t)$ 是非负定的.

必要性. 设 $f(t)$ 是非负定的, 则 (4.1) 式成立. 在 (4.1) 中取 $t_k = \frac{k}{n}, \xi_k = e^{-it_k x}(0 \leqslant k \leqslant N-1)$, 则

$$P_N^{(n)}(x) \overset{\text{def}}{=\!=} \frac{1}{N}\sum_{j=0}^{N-1}\sum_{k=0}^{N-1} f\left(\frac{k-j}{n}\right) e^{-i(k-j)x} \geqslant 0. \tag{4.6}$$

显然, (4.6) 右边的和中使 $k-j=r$ 的项数共有 $N-|r|$ 项 ($r$ 从 $-N+1$ 到 $N-1$). 所以

$$P_N^{(n)}(x) = \frac{1}{N}\sum_{r=-N+1}^{N-1}(N-|r|) f\left(\frac{r}{n}\right) e^{-irx}.$$

把上式两边乘以 $e^{isx}$ 并对 $x$ 由 $-\pi$ 到 $\pi$ 积分得:

$$\int_{-\pi}^{\pi} P_N^{(n)}(x)dx = \sum_{r=-N}^{N}\left(1-\frac{|r|}{N}\right) f\left(\frac{r}{n}\right)\int_{-\pi}^{\pi} e^{-i(r-s)x}dx. \tag{4.7}$$

由于

$$\int_{-\pi}^{\pi} e^{-i(r-s)x}dx = \begin{cases} 0, & s \neq r, \\ 2\pi, & s = r, \end{cases}$$

所以

$$\int_{-\pi}^{\pi} e^{-i(r-s)x}dx = \left(1-\frac{|s|}{N}\right) f\left(\frac{s}{n}\right) 2\pi$$

即

$$\left(1-\frac{|s|}{N}\right) f\left(\frac{s}{n}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(r-s)x}dx. \tag{4.8}$$

令

$$F_N^{(n)}(x) = \begin{cases} \frac{1}{2\pi}\int_{-\infty}^{x} P_N^{(n)}(y)dy, & \text{当 } -\pi \leqslant x \leqslant \pi, \\ F_N^{(n)}(\pi), & \text{当 } x > \pi, \\ F_N^{(n)}(-\pi), & \text{当 } x < -\pi, \end{cases}$$

则对任何固定的 $n$ 和 $N, F_N^{(n)}(x)$ 是 d.f., 根据第三章定理 5.6, $\{F_N^{(n)}(x), N = 1, 2, \cdots\}$ 有局部弱收敛子序列 $\{F_{N_k}^{(n)}(x), k \geqslant 1\}$. 但是

$$F_{N_k}^{(n)}(x) = \begin{cases} 0, & \text{当 } x \leqslant -\pi, \\ 1, & \text{当 } x > \pi, \end{cases}$$

所以存在分布函数 $F^{(n)}(x)$, 使

$$F_{N_k}^{(n)}(x) \xrightarrow{w} F^{(n)}(x) \quad (\text{当 } k \to \infty).$$

再用第三章定理 5.9, 有

$$\begin{aligned}&\lim_{k\to\infty}\int_{(-\infty,\pi]}e^{itx}F_{N_k}^{(n)}(dx)\\=&\lim_{k\to\infty}\int_{\mathbf{R}}e^{itx}F_{N_k}^{(n)}(dx)\\=&\int_{\mathbf{R}}e^{itx}F^{(n)}(dx)=\int_{[-\pi,\pi]}e^{itx}F^{(n)}(dx).\end{aligned}$$

因此

$$\begin{aligned}f\left(\frac{s}{n}\right)&=\lim_{k\to\infty}\left(1-\frac{|s|}{N}\right)f\left(\frac{s}{n}\right)\\&=\lim_{k\to\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}P_{N_k}^{(n)}(x)e^{isx}dx\\&=\int_{[-\pi,\pi]}e^{isx}F^{(n)}(dx).\end{aligned}\tag{4.9}$$

令

$$F_n(x)=F^{(n)}\left(\frac{x}{n}\right),$$

由于

$$F_n(x)=\begin{cases}1, & x>\pi,\\0 & x\leqslant-\pi,\end{cases}$$

所以 $F_n(x)$ 的 c.f. 为

$$f_n(t)=\int_{[-n\pi,n\pi]}e^{itx}F_n(dx),$$

从而

$$f_n\left(\frac{k}{n}\right)=\int_{[-\pi,\pi]}e^{iky}F_n(dy)=f\left(\frac{k}{n}\right).\tag{4.10}$$

对任意的 $t$, 选取 $k=k(n,t)$, 使 $0\leqslant t-\dfrac{k}{n}\leqslant\dfrac{1}{n}$, 则由 $f(t)$ 的连续性可知

$$f(t)=\lim_{n\to\infty}f\left(\frac{k}{n}\right)=\lim_{n\to\infty}f_n\left(\frac{k}{n}\right).\tag{4.11}$$

若能证 $\lim\limits_{n\to\infty}f_n(t)=f(t)$, 则由 $f_n(t)$ 是特征函数且 $f(t)$ 连续可知 $f(t)$ 也是特征函数. 定理获证.

事实上, 由 (4.10)、(4.11) 得

$$\begin{aligned}\lim_{n\to\infty}f_n(t)&=\lim_{n\to\infty}\left(\left[f_n(t)-f_n\left(\frac{k}{n}\right)\right]+f_n\left(\frac{k}{n}\right)\right)\\&=f(t)+\lim_{n\to\infty}\left(f_n(t)-f_n\left(\frac{k}{n}\right)\right).\end{aligned}\tag{4.12}$$

令 $\theta = t - \frac{k}{n}$, 按 $k$ 的取法有 $0 \leqslant \theta \leqslant \frac{1}{n}$, 而按函数 $f_n(t)$ 的定义有

$$\begin{aligned}\left|f_n(t) - f_n\left(\frac{k}{n}\right)\right| &= \left|\int_{[-n\pi,n\pi]} e^{i\frac{k}{n}x}\left(e^{i\theta x} - 1\right)F_n(dx)\right| \\ &\leqslant \int_{[-n\pi,n\pi]} \left|e^{i\theta x} - 1\right| F_n(dx). \end{aligned} \tag{4.13}$$

用 Hölder 不等式有

$$\begin{aligned}\int_{[-n\pi,n\pi]} \left|e^{i\theta x} - 1\right| F_n(dx) &\leqslant \sqrt{\int_{[-n\pi,n\pi]} \left|e^{i\theta x} - 1\right|^2 F_n(dx)} \\ &= \sqrt{\int_{[-n\pi,n\pi]} \left(e^{i\theta x} - 1\right)\left(e^{-i\theta x} - 1\right) F_n(dx)} \\ &= \sqrt{\int_{[-n\pi,n\pi]} 2(1 - \cos\theta x)F_n(dx)}. \\ &= \sqrt{2\mathscr{R}(1 - f_n(\theta))}. \end{aligned} \tag{4.14}$$

因为当 $0 \leqslant \varphi < 1, -\pi \leqslant z \leqslant \pi$ 时, $\cos z \leqslant \cos\varphi z$, 所以由 $0 \leqslant n\theta < 1$ 得

$$\begin{aligned}\mathscr{R}(1 - f_n(\theta)) &= \int_{[-n\pi,n\pi]} (1 - \cos\theta x)F_n(dx) \\ &\leqslant \int_{[-\pi,\pi]} (1 - \cos z)F^{(n)}(dz) \\ &= 1 - \mathscr{R}\left(\int_{[-\pi,\pi]} e^{iz}F^{(n)}(dz)\right) \\ &= 1 - \mathscr{R}\left(f\left(\frac{1}{n}\right)\right). \end{aligned} \tag{4.15}$$

综合 (4.13)、(4.14)、(4.15) 得

$$\left|f_n(t) - f_n\left(\frac{k}{n}\right)\right| \leqslant \sqrt{2\left[1 - \mathscr{R}\left(f\left(\frac{1}{n}\right)\right)\right]}. \tag{4.16}$$

由于 $f(t)$ 连续, $f(0) = 1$, 所以在 (4.16) 中令 $n \to \infty$ 取极限即得:

$$\lim_{n\to\infty}\left|f_n(t) - f_n\left(\frac{k}{n}\right)\right| = 0. \tag{4.17}$$

以 (4.17) 代入 (4.12) 得 $\lim\limits_{n\to\infty} f_n(t) = f(t)$. 定理 4.1 得证.

## *§5 随机元的特征泛函

在这一节中, 恒设 $(\Omega, \mathscr{F}, P)$ 为完备概率空间, $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \|\bullet\|)$ 为可测 Banach 空间, $X$ 是 $\boldsymbol{B}$ 值随机元, $P_X$ 是 $X$ 的分布, 即是 $\mathscr{B}(\boldsymbol{B})$ 上的概率测度满足 $P_X(A) = P(X \in A) = P(X^{-1}(A))$ $(A \in \mathscr{B}(\boldsymbol{B}))$. 有时记 $P_X = P \circ X^{-1}$. $\boldsymbol{B}^*$ 是 $\boldsymbol{B}$ 的共轭空间, 即 $\boldsymbol{B}$ 上的有界线性泛函全体.

**定义 5.1** 设 $X$ 是 $\boldsymbol{B}$ 值随机元, 其特征泛函定义为

$$f_X(l) = \int_{\boldsymbol{B}} e^{il(x)} P_X(dx) \quad (\forall l \in \boldsymbol{B}^*). \tag{5.1}$$

**定理 5.1** $\boldsymbol{B}$ 值随机元 $X$ 的特征泛函 $f_X(l)$ 有下述性质:

(1) $f_X(0) = 1$;

(2) $f_X(l)$ 是非负定的, 即对任何正整数 $n$, 任何 $l_1, l_2, \cdots, l_n \in \boldsymbol{B}^*$, 任何复数 $\xi_1, \xi_2, \cdots, \xi_n$, 都有

$$\sum_{j=1}^{n} \sum_{k=1}^{n} f_X(l_j - l_k) \xi_j \overline{\xi}_k \geqslant 0; \tag{5.2}$$

(3) $|f_X(l)| \leqslant 1, f_X(-l) = \overline{f_X(l)}$ $(\forall l \in \boldsymbol{B}^*)$;

(4) $f_X$ 在 $\boldsymbol{B}^*$ 上一致连续 (关于 $\boldsymbol{B}^*$ 中的强拓扑, 即关于算子范数形成的拓扑);

(5) 固定任意 $l_1, \cdots, l_d \in \boldsymbol{B}^*$, 给定 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), P_X)$ 上的取值于 $\mathbf{R}^d$ 的随机变量

$$V(x) = (l_1(x), \cdots, l_d(x)),$$

其特征函数为

$$f_V(t_1, \cdots, t_d) = f_X(t_1 l_1 + \cdots + t_d l_d).$$

(6) 特别地, 若 $\boldsymbol{B} = \mathbf{R}^d$, 则 $\boldsymbol{B}^*$ 中任一元素 $l$, 必存在 $t \in \mathbf{R}^d$, 使 $l(x) = \langle t, x \rangle, (\forall x \in \mathbf{R}^d)$, 此处 $\langle t, x \rangle$ 是 $t$ 与 $x$ 之内积. 反之也对. 于是, 取值于 $\mathbf{R}^d$ 的随机向量的特征函数为

$$f_X(t_1, \cdots, t_d) = \int_{\mathbf{R}^d} e^{i \sum\limits_{k=1}^{d} t_k x_k} P_X(dx_1, \cdots, dx_d)((t_1, \cdots, t_d) \in \mathbf{R}^d), \tag{5.3}$$

这与通常的定义是一致的.

**证** 只证 (4), 其余诸结论是显然的. $\boldsymbol{B}^*$ 中的范数用 $\|\bullet\|^*$ 表示. 任取 $l, l' \in \boldsymbol{B}^*$, 显然有:

$$|f_X(l) - f_X(l')| \leqslant \int_{\boldsymbol{B}} \left| e^{il(x)} - e^{il'(x)} \right| P_X(dx)$$

$$\leqslant \int_{\boldsymbol{B}} \left| e^{i(l(x) - l'(x))} - 1 \right| P_X(dx). \tag{5.4}$$

$\forall\varepsilon > 0$, 存在 $M > 0$, 使 $P_X(\{x \in \boldsymbol{B} : \|x\| \geqslant M\}) < \dfrac{\varepsilon}{4}$. 再取 $\delta > 0$ 充分小, 使 $|t| < \delta$ 时, $|e^{it} - 1| < \dfrac{\varepsilon}{2}$. 则当 $\|l - l'\|^* < \dfrac{\delta}{M}$ 时, 有

$$\begin{aligned}\int_{\boldsymbol{B}} \left|e^{i(l(x)-l'(x))} - 1\right| P_X(dx) &\leqslant \int_{\{\|x\|\leqslant M\}} \left|e^{i(l(x)-l'(x))} - 1\right| P_X(dx) \\ &\quad + 2P_X(\|x\| > M) \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \tag{5.5}$$

由 (5.4)、(5.5) 可得结论 (4).

**定理 5.2** 设 $(\boldsymbol{B}, \|\bullet\|)$ 是可分的 Banach 空间, $\mu_1$ 和 $\mu_2$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的两个概率测度, 则 $\mu_1 \equiv \mu_2$ 的充分必要条件是:

$$\int_{\boldsymbol{B}} e^{il(x)}\mu_1(dx) = \int_{\boldsymbol{B}} e^{il(x)}\mu_2(dx) \quad (\forall l \in \boldsymbol{B}^*). \tag{5.6}$$

**证** 只证充分性, 必要性是显然的. 事实上, 任取 $l_1, \cdots, l_d \in \boldsymbol{B}^*, t_1, \cdots, t_d \in \mathbf{R}$, 令

$$\varphi_j(t_1, \cdots, t_d) = \int_{\boldsymbol{B}} e^{i\sum\limits_{k=1}^{d} t_k l_k(x)} P_X(dx) \quad (j = 1, 2), \tag{5.7}$$

则由 (5.6) 和 (5.7) 可知:

$$\varphi_1(t_1, \cdots, t_d) = \varphi_2(t_1, \cdots, t_d). \tag{5.8}$$

但是 $\varphi_j(t_1, \cdots, t_d)$ 是定义在 $(\boldsymbol{B}, \mathscr{B}(\boldsymbol{B}), \mu_j)$ 上取值于 $\mathbf{R}^d$ 的随机向量

$$V_j(x) = (l_1(x), \cdots, l_d(x))$$

的特征函数 $(j = 1, 2)$. 既然 $\varphi_1 \equiv \varphi_2$, 所以 $V_1(x)$ 的分布函数与 $V_2(x)$ 的分布函数恒等. (参见 [36] 第二章定理 7.2.) 亦即对任何 $A_d \in \mathscr{B}(\mathbf{R}^d)$,

$$\begin{aligned}&\mu_1\left(\{x \in \boldsymbol{B} : (l_1(x), \cdots, l_d(x)) \in A_d\}\right) \\ &= \mu_2\left(\{x \in \boldsymbol{B} : (l_1(x), \cdots, l_d(x)) \in A_d\}\right). \end{aligned} \tag{5.9}$$

注意 (5.9) 是对任意正整数 $d$, 任意 $A_d \in \mathscr{B}(\mathbf{R}^d)$, 任意 $l_1, \cdots, l_d \in \boldsymbol{B}^*$ 成立的, 即 $\mu_1$ 与 $\mu_2$ 在

$$\mathscr{E} = \{\{x \in \boldsymbol{B} : (l_1(x), \cdots, l_d(x)) \in A_d\} : A_d \in \mathscr{B}(\mathbf{R}^d), l_1, \cdots, l_d \in \boldsymbol{B}^*, d \geqslant 1\}$$

上恒等. 但是 $\mathscr{E}$ 是一个代数, 且它产生的 $\sigma$ 代数就是 $\sigma(\mathscr{E}) = \mathscr{B}(\boldsymbol{B})$. 所以由测度扩张之唯一性知:

$$\mu_1(A) = \mu_2(A) \quad (\forall A \in \mathscr{B}(\boldsymbol{B})).$$

**定理 5.3** 设 $(\boldsymbol{B},\|\bullet\|)$ 是可分的 Banach 空间, $\mu,\mu_1,\mu_2$ 是 $\mathscr{B}(\boldsymbol{B})$ 上的概率测度, 则 $\mu=\mu_1 * \mu_2$ 的充分必要条件是:

$$\int_{\boldsymbol{B}} e^{il(x)}\mu(dx)=\int_{\boldsymbol{B}} e^{il(x)}\mu_1(dx)\bullet\int_{\boldsymbol{B}} e^{il(x)}\mu_2(dx),\quad \forall l\in \boldsymbol{B}^*, \tag{5.10}$$

即是说概率测度的卷积的特征泛函是各自的特征泛函的乘积.

**证** 必要性. 令

$$T:\boldsymbol{B}\times\boldsymbol{B}\mapsto\boldsymbol{B},\quad T(x,y)=x+y,\quad x,y\in\boldsymbol{B}.$$

若 $\mu=\mu_1 * \mu_2$, 则 $\mu=(\mu_1\times\mu_2)\circ T^{-1}$. 所以对任何 $l\in\boldsymbol{B}^*$, 有

$$\begin{aligned}\int_{\boldsymbol{B}} e^{il(x)}\mu(dx)&=\int_{\boldsymbol{B}} e^{il(x)}((\mu_1\times\mu_2)\circ T^{-1})(dx)\\&=\int_{\boldsymbol{B}\times\boldsymbol{B}} e^{il(x+y)}\mu_1(dx)\mu_2(dy)\\&=\int_{\boldsymbol{B}} e^{il(x)}\mu_1(dx)\bullet\int_{\boldsymbol{B}} e^{il(y)}\mu_2(dy).\end{aligned}\tag{5.11}$$

充分性. 由于

$$\int_{\boldsymbol{B}} e^{il(x)}(\mu_1 * \mu_2)(dx)=\int_{\boldsymbol{B}} e^{il(x)}\mu_1(dx)\bullet\int_{\boldsymbol{B}} e^{il(y)}\mu_2(dy),$$

再用定理 5.2 可知充分性成立.

**定理 5.4** 设 $(\boldsymbol{B},\|\bullet\|)$ 是可分的 Banach 空间, $X_1$ 与 $X_2$ 是相互独立的 $\boldsymbol{B}$ 值随机元 (从而 $X_1+X_2$ 也是 $\boldsymbol{B}$ 值元), 则 $X_1+X_2$ 的特征泛函是 $X_1$ 的特征泛函与 $X_2$ 的特征泛函的乘积, $X_1+X_2$ 的分布是 $X_1$ 的分布与 $X_2$ 的分布的卷积.

**证** 由定理 5.2 和 5.3 立即可得定理 5.4.

**定理 5.5** 设 $\{X,X^{(n)},n\geqslant 1\}$ 是一列 $\boldsymbol{B}$ 值随机元 (并不要求 $(\boldsymbol{B},\|\bullet\|)$ 是可分的), $\{P_X,P_{X^{(n)}},n\geqslant 1\}$ 是相应的分布列, $\{f_X,f_{X^{(n)}},n\geqslant 1\}$ 是相应的特征泛函列, 则

$$P_{X^{(n)}}\xrightarrow{\text{w}}P_X \text{ 蕴涵了 }\quad f_{X^{(n)}}(l)\to f_X(l)\quad(\forall l\in\boldsymbol{B}^*).$$

**证** 设 $P_{X^{(n)}}\xrightarrow{\text{w}}P_X$. 由于对任何固定的 $l\in\boldsymbol{B}^*, e^{il(x)}$ 是定义在 $\boldsymbol{B}$ 上的取值于复空间的有界连续函数, 所以由弱收敛的定义有

$$\begin{aligned}f_{X^{(n)}}(l)=\int_{\boldsymbol{B}} e^{il(x)}P_{X^{(n)}}(dx)&\to\int_{\boldsymbol{B}} e^{il(x)}P_X(dx)\\&=f_X(l).\end{aligned}$$

**附注 5.1** 本定理之逆命题未必成立 (参见 §6 习题 7). 但本定理之逆命题对 $\boldsymbol{B}=\mathbf{R}^d$ 这一特殊情形是对的. 这说明一般的 Banach 空间 $(\boldsymbol{B},\|\bullet\|)$ 可能是无穷维的, 对有穷维空间 $\mathbf{R}^d$ 成立的命题对无穷维空间未必成立. 因此, 对 $\mathbf{R}^d$ 空间中有而对一般的 Banach 空间无的命题, 还得要提一下.

**定理 5.6** 设 $X$ 是取值于 $\mathbf{R}^d$ 的随机向量, $P_X$ 和 $f_X$ 分别是 $X$ 的分布和特征函数, 则对 $P_X$ 的任一连续区间 $(a,b]=\{(x_1,\cdots,x_d)\in\mathbf{R}^d: a_i<x_i\leqslant b_i, i=1,\cdots,d\}, u=(a_1,\cdots,a_d), b=(b_1,\cdots,b_d)$, 均有:

$$\begin{aligned}P_X((a,b))&=P_X((a,b])=P_X([a,b])\\&=\lim_{\substack{t_k\to\infty\\1\leqslant k\leqslant d}}\int_{-t_1}^{t_1}\cdots\int_{-t_d}^{t_d}\left(\prod_{k=1}^{d}\frac{e^{-it_ka_k}-e^{-it_kb_k}}{i2\pi t_k}\right)\bullet\\&\quad f_X(t_1,\cdots,t_d)dt_1\cdots dt_d,\end{aligned}$$

证明请参见 [36] 第二章定理 7.2.

**定理 5.7** 设 $\{X,X^{(n)},n\geqslant 1\}$ 是一列取值于 $\mathbf{R}^d$ 的随机向量, 其对应的分布列为 $\{P_X,P_{X^{(n)}},n\geqslant 1\}$, 对应的特征函数列是 $\{f_X,f_{X^{(n)}},n\geqslant 1\}$, 则

$$P_{X^{(n)}}\overset{\mathrm{w}}{\Longrightarrow}P_X\Leftrightarrow f_{X^{(n)}}\to f_X.$$

证明请参见 [36] 第二章定理 7.6 和定理 7.7.

## §6 习题及应用

1. 试证下列诸对应关系:

(a) 二项分布, 其分布为 $p_k=\binom{n}{k}p^k(1-p)^{n-k}(k=0,1,\cdots,n)$, 其特征函数为 $f(t)=(pe^{it}+(1-p))^n$;

(b) 参数为 $\lambda$ 的 Poisson 分布, 其分布为 $p_k=e^{-\lambda}\dfrac{\lambda^k}{k!}(k=0,1,2,\cdots)$, 其特征函数为 $f(t)=e^{-\lambda(e^{it}-1)}$;

(c) 均匀分布, 其密度函数为

$$p(x)=\begin{cases}\dfrac{1}{b-a}, & \text{当 } a\leqslant x\leqslant b,\\ 0, & \text{反之,}\end{cases}$$

其特征函数为 $\left(e^{ibt}-e^{iat}\right)/i(b-a)t$;

(d) Cauchy 分布, 其密度函数为

$$p(x)=\frac{1}{\pi}\frac{a}{a^2+(x-b)^2}\quad (a>0),$$

其特征函数为

$$f(t)=e^{-a|t|+ibt};$$

(e) Laplace 分布, 其密度函数为

$$p(x)=\frac{1}{2a}e^{-|x-b|/a} \quad (a>0,-\infty<b<\infty),$$

其特征函数为

$$f(t)=e^{ibt}(1+a^2t^2)^{-1}.$$

(f) 正态分布, 其密度函数为

$$p(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-a}{\sigma}\right)^2} \quad (\sigma>0),$$

其特征函数为

$$f(t)=e^{iat-\frac{1}{2}\sigma^2t^2}.$$

2. 试证下列诸函数是特征函数, 并求出其对应的分布函数.

(a) $f_1(t)=\sum\limits_{k=0}^{\infty}a_k\cos kt \quad (a_k\geqslant 0)$;

(b) $f_2(t)=\sum\limits_{k=0}^{\infty}a_ke^{i\lambda_kt} \quad (a_k\geqslant 0)$;

(c) $f_3(t)=\cos^2 t$;

(d) $f_4(t)=(\sin\ at)/at$.

3. 若 $f(t)$ 是随机变量 $X$ 的特征函数, 则下列函数也是特征函数:

(a) $f_1(t)=e^{f(t)-1}$;

(b) $f_2(t)=\dfrac{1}{t}\int_0^t f(s)ds$ (此处 $\dfrac{0}{0}$ 定义为 1).

4. 设 $\{F(x),F_k(x),k\geqslant 1\}$ 为一列 d.f., 其对应的 c.f. 列为 $\{f(t),f_k(t),k\geqslant 1\}$. 令

$$M_hf=\frac{2}{\pi h}\int_0^{\infty}|f(t)|^2\frac{\sin^2 ht}{t}dt \quad (h>0),$$

$$Mf=\lim_{T\to\infty}\int_{-T}^{T}|f(t)|^2dt,$$

则

(a) $M_hf$ 是 $h$ 的单调非降函数, 而且

$$\lim_{h\to\infty}M_hf=1, \quad \lim_{h\to 0}M_hf=Mf;$$

(b) $\lim\limits_{m\to\infty}\lim\limits_{n\to\infty}M_h\left(\prod\limits_{k=m+1}^{n}f_k\right)$

或则为 0 或则为 1;

(c) $M(f)=\sum\limits_k(F(x_k)-F(x_k-0))^2$, 其中 $\{x_k\}$ 为 $F(x)$ 的全部间断点;

(d) $M(f_1f_2)\geqslant M(f_1)M(f_2)$

$$M_hf=2\int_0^{2h}\left(1-\frac{x}{2h}\right)F^s(dx),$$

其中 $F^s(x)$ 是 $|f(t)|^2$ 所对应的 d.f.;

(e) $\lim\limits_{n\to\infty}f_n(t)=f(t)$ 且 $Mf=0\Rightarrow\lim\limits_{n\to\infty}Mf_n=0$, 但本结论之逆不一定成立;

(f) $\lim\limits_{n\to\infty}\prod\limits_{k=1}^n f_k(t)=f(t)\Rightarrow\lim\limits_{n\to\infty}M\left(\prod\limits_{k=1}^n f_k\right)=M(f)$.

5. 若 d.f.$F(x)$ 有密度函数, 则 $F(x)$ 对应之 c.f.$f(t)$ 在 $|t|\to\infty$ 时趋于 0.

6. 若 R.V.$X$ 满足:

$$\lim_{x\to\infty}x^nP(|X|\geqslant x)=0,\quad 对一切\ n\geqslant\ 上成立,$$

则 $X$ 的任意阶矩都存在.

7. 设 R.V.$X_1$ 和 $X_2$ 对应的 c.f. 分别为 $f_1(t)$ 和 $f_2(t)$, 随机向量 $(X_1,X_2)$ 的 c.f. 为 $f(t_1,t_2)$, 则 $X_1$ 与 $X_2$ 相互独立的充分必要条件是 $f(t_1,t_2)=f_1(t_1)f_2(t_2)$.

8. 设 $\{\boldsymbol{e}_k,k\geqslant 1\}$ 是无穷维的可分 Hilbert 空间 $\boldsymbol{H}$ 中的一组标准正交基, $\mu_k$ 是在 $\boldsymbol{e}_k$ 上的退化分布, 即 $\mu_k(\{\boldsymbol{e}_k\})=1,\mu_k(\boldsymbol{H}-\{\boldsymbol{e}_k\})=0$. 再令 $\delta_0$ 是在 $O$ 点的退化分布. 注意: 现在 $\boldsymbol{H}$ 是 Hilbert 空间, 故有 $\boldsymbol{H}^*=\boldsymbol{H}$. 令 $\mu_k$ 对应的特征泛函为

$$\begin{aligned}f_k(l)&=\int_{\boldsymbol{H}}e^{il(x)}\mu_k(dx)=\int_{\boldsymbol{H}}e^{i\langle l,x\rangle}\mu_k(dx)\\&=e^{i\langle l,\boldsymbol{e}_k\rangle},\quad (l\in\boldsymbol{H}^*=\boldsymbol{H},\langle x,y\rangle\ 表\ x\ 和\ y\ 之内积.)\ 证明:\end{aligned}$$

(1) $\lim\limits_{k\to\infty}f_k(l)\equiv 1,(\forall l\in\boldsymbol{H}^*=\boldsymbol{H})$;

(2) $\delta_0$ 的特征泛函 $f_0(l)\equiv 1(\forall l\in\boldsymbol{H}^*=\boldsymbol{H})$;

(3) 但是绝不可能 $\mu_k\overset{\mathrm{w}}{\Longrightarrow}\delta_0$, 当 $k\to\infty$ 时. (提示: 取有界连续函数 $g:\boldsymbol{H}\to\mathbf{R}$ 满足 $g(0)=1,\ g(x)=0$ (当 $\|x\|\geqslant\dfrac{1}{2}$).)

9. 证明定理 2.5 中的增量不等式、积分不等式、截尾不等式及几个积分的估计.

# 第五章　大数定律、中心极限定理、重对数律

在这一章中, 我们研究经典的概率极限理论中的基础部分, 即 (弱) 大数定律、强大数定理、中心极限定理和重对数律.

如不特别声明, 本章中恒用 R.V. 表示实值随机变量, d.f 和 c.f. 分别表示分布函数与特征函数, $\delta_a$ 表示仅在单点集 $\{a\}$ 上有测度 1 而在其他集上测度为 0 的退化分布, 或者表示仅在点 $a$ 上有跃度 1 的阶梯分布函数, 特别地, 称 $\delta_0$ 为 $0-1$ 律. 如不声明, 随机变量总是定义在完备概率空间 $(\Omega, \mathscr{F}, P)$ 上的。

## §1　独立同分布随机变量列的大数定律

**定义 1.1**　设 $\{X_k, k \geqslant 1\}$ 是 $(\Omega, \mathscr{F}, P)$ 上的随机变量列, 如果存在实数列 $\{a_n, n \geqslant 1\}$, 使得对任何 $\varepsilon > 0$, 都有

$$\lim_{n \to \infty} P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - a_n\right| \geqslant \varepsilon\right) = 0, \tag{1.1}$$

即

$$\left(\frac{1}{n}\sum_{k=1}^{n} X_k - a_n\right) \xrightarrow{P} 0,$$

则称 $\{X_k, k \geqslant 1\}$ 服从弱大数定律, 简称服从大数定律.

**命题 1.1** 设随机变量列 $\{X, X_n, n \geqslant 1\}$ 对应的分布函数列为 $\{F(x), F_n(x), n \geqslant 1\}$, 则

$$X_n \xrightarrow{P} X \Longrightarrow F_n(x) \xrightarrow{w} F(x),$$

但逆命题不对.

**证** 任取函数 $F(x)$ 的连续点 $x$, 都有

$$\begin{aligned}
&|F_n(x) - F(x)| \leqslant |P(X_n \leqslant x) - P(X \leqslant x+\varepsilon)| \\
&\quad +|P(X \leqslant x+\varepsilon) - P(X \leqslant x)| \\
&\leqslant P(X_n \leqslant x, X > x+\varepsilon) + P(X_n > x, X \leqslant x+\varepsilon) \\
&\quad +|F(x+\varepsilon) - F(x)| \\
&\leqslant P(X_n \leqslant x, X > x+\varepsilon) + P(X_n > x, X \leqslant x-\varepsilon) \\
&\quad +P(x-\varepsilon < X \leqslant x+\varepsilon) + |F(x+\varepsilon) - F(x)| \\
&\leqslant 2P(|X_n - X| \geqslant \varepsilon) + F(x+\varepsilon) - F(x-\varepsilon) \\
&\quad +F(x+\varepsilon) - F(x).
\end{aligned} \tag{1.2}$$

由 $X_n \xrightarrow{P} X$ 及 $x$ 是 $F(x)$ 的连续点可知: 在 (1.2) 上令 $\varepsilon \to 0$ 可得

$$F_n(x) \to F(x).$$

所以 $F_n(x) \xrightarrow{w} F(x)$.

反例如下: 作特殊的概率空间如下: $\Omega = \{0,1\}$, $\mathscr{F} = \{\phi, \Omega, \{0\}, \{1\}\}$, $P(\{0\}) = P(\{1\}) = \dfrac{1}{2}$. 再定义随机变量列 $\{X_n, n \geqslant 1\}$ 如下:

$$X_{2n+1}(\omega) = \begin{cases} 0, & \text{当 } \omega = 0, \\ 1, & \text{当 } \omega = 1, \end{cases} \quad n = 0, 1, 2, \cdots,$$

$$X_{2n}(\omega) = \begin{cases} 0, & \text{当 } \omega = 1, \\ 1, & \text{当 } \omega = 0, \end{cases} \quad n = 1, 2, \cdots,$$

则 $X_n$ 的分布函数 $F_n(x)$ 为

$$F_n(x) = \begin{cases} 0, & \text{当 } x < 0, \\ \dfrac{1}{2}, & \text{当 } 0 \leqslant x < 1, \\ 1, & \text{当 } x \geqslant 1, \end{cases} \quad n = 1, 2, \cdots,$$

所以 $\{F_n(x), n \geqslant 1\}$ 弱收敛, 但 $\{X_n, n \geqslant 1\}$ 绝不能依概率收敛.

尽管命题 1.1 说, 一般地随机变量列的依概率收敛并不等价于依分布收敛, 但在特殊场合, 它们是等价的. 请看

**命题 1.2** 设 R.V. $X_n$ 的分布函数为 $F_n(x)(n \geqslant 1)$, $a$ 为一个常数, $\delta_a$ 是在 $a$ 点的退化分布, 则

$$X_n \xrightarrow{P} a \Leftrightarrow F_n(x) \xrightarrow{\mathrm{w}} \delta_a.$$

**证** 必要性是命题 1.1 的特款, 故只需证明充分性. 设

$$F_n(x) \xrightarrow{\mathrm{w}} \delta_a. \tag{1.3}$$

因为对任何 $\varepsilon > 0$,

$$P(|X_n - a| < \varepsilon) \leqslant F_n(a+\varepsilon) - F_n(a-\varepsilon). \tag{1.4}$$

而 $(a+\varepsilon)$ 和 $(a-\varepsilon)$ 都是退化分布函数 $\delta_a$ 的连续点, 所以, 由 (1.3) 和 (1.4) 可得

$$\lim_{n \longrightarrow \infty} P(|X_n - a| < \varepsilon) = 1 \quad (\forall \varepsilon > 0),$$

此即 $X_n \xrightarrow{P} a$. 命题 1.2 得证.

根据命题 1.2 和第四章的推论 2.3, 若要使随机变量列 $\{X_k, k \geqslant 1\}$ 服从大数定律, 必须而且只需存在一列实数 $\{a_n, n \geqslant 1\}$ 使下面任何一个条件成立:

1. $\dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - a_k) \xrightarrow{P} 0$;
2. $P \circ \left(\dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - a_k)\right)^{-1} \xRightarrow{\mathrm{w}} \delta_0$;
3. $\boldsymbol{E}\left(\exp\left(it\dfrac{1}{n}\sum\limits_{k=1}^{n}(X_k - a_k)\right)\right) \longrightarrow 1$. 此处 $\exp(x)$ 表 $e^x$.

**定理 1.1 (Khinchin)** 设 $\{X_n, n \geqslant 1\}$ 是相互独立、具有公共分布函数 $F(x)$ 的随机变量序列, $f(t)$ 是其公共的特征函数, 如果 $\boldsymbol{E}(X_n) = 0$ 存在, 令

$$S_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \quad n = 1, 2, \cdots,$$

则

$$\boldsymbol{E}(\exp(itS_n/n)) \to 1,$$

即 $\{X_n, n \geqslant 1\}$ 服从大数定律.

**证** 由于 $\{X_n, n \geqslant 1\}$ 相互独立且有公共的特征函数 $f(t)$, 所以

$$\begin{aligned}&\boldsymbol{E}(\exp(itS_n/n)) \\ =& \prod_{k=1}^{n} \boldsymbol{E}(\exp(itX_k/n)) = f\left(\frac{t}{n}\right)^n,\end{aligned}$$

但是

$$f\left(\frac{t}{n}\right) = f(0) + f'(0)\frac{t}{n} + o\left(\frac{t}{n}\right)$$
$$= 1 + o\left(\frac{t}{n}\right) \quad (n \to \infty).$$

所以

$$\boldsymbol{E}(\exp(itS_n/n)) = \left(1 + o\left(\frac{t}{n}\right)\right)^n \to 1.$$

定理证毕.

**推论 1.1** 若 $\{X_n, n \geqslant 1\}$ 是相互独立具有公共分布函数的随机变量序列, 且 $\boldsymbol{E}(X_n) = p$ 存在, 则

$$\lim_{n\to\infty} \boldsymbol{E}(\exp(it(S_n - np)/n)) = 1,$$

即是 $\{X_n, n \geqslant 1\}$ 也服从大数定律.

**例 1.1 (Benoulli 大数定律)** 设 $S_n$ 是 $n$ 次相互独立的试验中事件 $A$ 出现的次数, $p$ 是事件 $A$ 在每次试验中出现的概率, 则由推论 1.1 及命题 1.2 有

$$\frac{S_n}{n} \xrightarrow{P} p,$$

此即 "频率 $S_n/n$ 趋于概率 $p$".

由此可见, 大数定律给概率的统计定义提供了一个理论根据. 历史上将例 1.1 称为 Bernoulli 大数定律, 也是最早、最简单的一个大数定律.

## §2 独立同分布随机变量列的中心极限定理

**定义 2.1** 设 $\{X_n, n \geqslant 1\}$ 是随机变量列, 如果存在一列实数 $\{a_n, n \geqslant 1\}$ 和一列正数 $\{b_n, n \geqslant 1\}$, 使

$$P \circ \left(\frac{\sum_{k=1}^{n}(X_k - a_k)}{\sum_{k=1}^{n} b_k}\right)^{-1} \xrightarrow{\mathrm{w}} N(0, 1),$$

则称 $\{X_n\}$ 服从中心极限定理, 其中 $N(a, \sigma^2)$ 表示数学期望为 $a$, 方差为 $\sigma^2$ 的正态分布.

**定理 2.1** 设 $\{X_n, n \geqslant 1\}$ 是一列相互独立且具有公共分布的随机变量, 其公共方差 $\mathrm{var}(X_n) = \sigma_n^2 = \sigma^2$ 存在, 且 $\boldsymbol{E}(X_n) = 0$. 令

$$S_n = \sum_{k=1}^{n} X_k,\ s_n^2 = \mathrm{var}(S_n) = n\sigma^2,$$

则

$$P \circ \left(\frac{S_n}{s_n}\right)^{-1} \xrightarrow{\mathrm{w}} N(0,1),$$

即 $\{X_n\}$ 服从中心极限定理.

**证** 令 $f(t)$ 为 $\{X_n\}$ 的公共特征函数, 由于 $X_n$ 的二阶矩存在, 且 $\boldsymbol{E}(X_n) = 0$, 所以

$$\begin{aligned} f\left(\frac{t}{s_n}\right) &= f(0) + f'(0)\frac{t}{s_n} + \frac{f''(0)}{2}\left(\frac{t}{s_n}\right)^2 + o\left(\left(\frac{t}{s_n}\right)^2\right) \\ &= 1 - \frac{1}{2}\left(\frac{t\sigma}{s_n}\right)^2 + o\left(\frac{t^2}{s_n^2}\right) \\ &= 1 - \frac{1}{n}\frac{t^2}{2} + o\left(\frac{t^2}{n}\right), \end{aligned}$$

从而

$$\boldsymbol{E}(\exp(itS_n/s_n)) = \left(1 - \frac{1}{n}\frac{t^2}{2} + o\left(\frac{t^2}{n}\right)\right)^n \to e^{-\frac{1}{2}t^2},$$

此即

$$P \circ \left(\frac{S_n}{s_n}\right)^{-1} \xrightarrow{\mathrm{w}} N(0,1).$$

定理证毕.

**推论 2.1** 若 $\{X_n, n \geqslant 1\}$ 为相互独立且具有公共分布的随机变量列, 且 $\boldsymbol{E}(X_n) = p, \mathrm{var}(X_n) = \sigma_n^2 = \sigma^2, 0 < \sigma^2 < \infty$, 则

$$P \circ \left(\frac{S_n - np}{s_n}\right)^{-1} \xrightarrow{\mathrm{w}} N(0,1),$$

其中 $S_n = \sum\limits_{k=1}^{n} X_k, s_n^2 = n\sigma^2$.

**例 2.1 (De Moivre – Laplace 中心极限定理)** 设 $S_n$ 为 $n$ 次独立试验中事件 $A$ 出现的次数, 而每次试验 $A$ 出现的概率为 $p \in (0,1)$, 则

$$P \circ \left(\frac{S_n - np}{\sqrt{np(1-p)}}\right)^{-1} \xrightarrow{\mathrm{w}} N(0,1).$$

这是最早也是最简单的一个中心极限定理.

## §3 独立随机变量列的大数定律

在第 1 节中, 我们证明了独立同分布的随机变量列, 只要它们的公共数学期望存在, 大数定律总是成立. 但在分布不同的场合, 情况就复杂一些. 下面我们给出对一般独立的随机变量列而言, 它满足大数定律的各种充分条件.

在正式讲大数定律以前, 我们先讲几点附注.

**附注 3.1** 特征函数的对数. 这本是多值函数, 但我们约定取主值.

若 c.f. $f(t)$ 在 $|t| < T$ 内无处为 0, 则 $\log f(t)$ 定义为满足下述二条件的连续函数 $\varphi(t)$:

(1) $e^{\varphi(t)} = f(t),\ |t| < T$;

(2) $\varphi(0) = 0$.

**附注 3.2** 关于两个分布函数列的等价性.

称两个分布函数列 $\{F_n(x), n \geqslant 1\}$ 和 $\{G_n(x), n \geqslant 1\}$ (或相应的 (L-S) 测度 $\{F_n, n \geqslant 1\}$ 和 $\{G_n, n \geqslant 1\}$) 是等价的, 如果对 $\{F_n(x), n \geqslant 1\}$ 的任何一个弱收敛子列 $\{F_{n_k}(x),\ k \geqslant 1\}$,

$$F_{n_k}(x) \xrightarrow{w} F(x) \quad (k \to \infty),$$

都有

$$G_{n_k}(x) \xrightarrow{w} F(x), \quad (k \to \infty),$$

且反之亦然.

由上述定义看出: 若 $F_n(x) \xrightarrow{w} F(x)$, 欲证 $G_n(x) \xrightarrow{w} F(x)$, 只需证明 $\{F_n(x), n \geqslant 1\}$ 与 $\{G_n(x), n \geqslant 1\}$ 等价就行了.

**命题 3.1 (等价性引理)** 若 $P(X_n \neq Y_n) \to 0$, 则 $\{P \circ X_n^{-1}, n \geqslant 1\}$ 与 $\{P \circ Y_n^{-1}, n \geqslant 1\}$ 等价.

**证** 由

$$\begin{aligned}|P(X_n \leqslant x) - P(Y_n \leqslant x)| &\leqslant P(X_n \leqslant x, Y_n > x) + P(Y_n \leqslant x, X_n > x) \\ &\leqslant 2P(X_n \neq Y_n)\end{aligned}$$

立即可得命题 3.1.

**附注 3.3** 关于中位数.

**命题 3.2** 若 R.V. $X$ 满足 $P(a \leqslant X \leqslant b) > \dfrac{1}{2}$, 则 $X$ 在 $[a, b]$ 上至少有一个中位数.

**证** 令 $\alpha=\sup\left\{x: x\in[a,b], P(X\leqslant x)<\frac{1}{2}\right\}$, 则 $\alpha$ 就是 $X$ 的一个中位数.

注意: 中位数恒存在, 但不一定唯一.

**命题 3.3** 若 R.V. $X$ 满足 $P(|X|\geqslant\varepsilon)<\frac{1}{2}$, $\alpha$ 是 $X$ 的中位数, 则 $|\alpha|<\varepsilon$.

**证** 若 $\alpha>\varepsilon$, 则

$$P(|X|\geqslant\varepsilon)\geqslant P(X\geqslant\varepsilon)\geqslant P(X\geqslant\alpha)\geqslant\frac{1}{2},$$

这与命题的假设矛盾.

若 $\alpha<-\varepsilon$, 则

$$P(|X|\geqslant\varepsilon)\geqslant P(X\leqslant-\varepsilon)\geqslant P(X\leqslant\alpha)\geqslant\frac{1}{2},$$

这也与命题的假设矛盾.

总之, $|\alpha|<\varepsilon$.

**命题 3.4** 设 $\alpha$ 是 R.V. $X$ 的一个中位数, $X'$ 与 $X$ 独立同分布. 定义 $X^s=X-X'$ 为 $X$ 的对称化 R.V. (注意 $P(X^s\leqslant x)\equiv P(-X^s\leqslant x)$), 则

$$\frac{1}{2}P(|X-\alpha|\geqslant x)\leqslant P(|X^s|\geqslant x).$$

**证**

$$\begin{aligned}
P(X^s\leqslant -x)&\geqslant P(X\leqslant\alpha-x, X'\geqslant\alpha)\\
&=P(X\leqslant\alpha-x)P(X'\geqslant\alpha)\geqslant\frac{1}{2}P(X\leqslant\alpha-x).\\
&=\frac{1}{2}P(X-\alpha\leqslant-x).\\
P(X^s\geqslant x)&\geqslant P(X\geqslant\alpha+x, X'\leqslant\alpha)\\
&=P(X\geqslant\alpha+x)P(X'\leqslant\alpha)\\
&\geqslant P(X-\alpha\geqslant x)\cdot\frac{1}{2}.
\end{aligned}$$

由上述二不等式即得命题 3.4.

**定理 3.1 (Chebyshev)** 设 $\{X_n, n\geqslant 1\}$ 是相互独立的满足 $\operatorname{var}(X_n)\leqslant C<\infty$ $(n\geqslant 1)$ 的随机变量序列. 令 $S_n=\sum\limits_{k=1}^{n}X_k$, 则

$$P\circ\left(\frac{S_n-\boldsymbol{E}(S_n)}{n}\right)^{-1}\stackrel{\mathrm{w}}{\Longrightarrow}\delta_0.$$

(其中 $\delta_0$ 是 $0-1$ 律. 今后不再说明了.)

**证** 利用第三章定理 2.1 的 Chebyshev 不等式可得

$$
\begin{aligned}
&P\left(\left|\frac{S_n-\boldsymbol{E}(S_n)}{n}\right|\geqslant\varepsilon\right)\\
\leqslant&\frac{1}{\varepsilon^2}\mathrm{var}\left(\frac{S_n-\boldsymbol{E}(S_n)}{n}\right)\leqslant\frac{1}{(n\varepsilon)^2}\mathrm{var}(S_n)\\
\leqslant&\frac{C}{n\varepsilon^2}.
\end{aligned}
$$

所以

$$P\circ\left(\frac{S_n-\boldsymbol{E}(S_n)}{n}\right)^{-1}\xrightarrow{\mathrm{w}}\delta_0.$$

**例 3.1 (Poisson 大数定律)** 设 $S_n$ 是 $n$ 次相互独立的试验中事件 $A$ 出现的次数, 而 $p_k$ 是事件 $A$ 在第 $k$ 次试验中出现的概率, 则

$$P\circ\left(\frac{S_n}{n}-\frac{p_1+\cdots+p_n}{n}\right)^{-1}\xrightarrow{\mathrm{w}}\delta_0.$$

**定理 3.2 (Markov)** 设 $\{X_n,n\geqslant1\}$ 是相互独立的随机变量序列, $\boldsymbol{E}(X_n)=0$ $(n=1,2,\cdots)$, 且对某个 $0<\delta<1$ 有

$$\frac{1}{n^{1+\delta}}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^{1+\delta}\right)\to0\quad(n\to\infty),$$

则

$$P\circ\left(\frac{S_n}{n}\right)^{-1}\xrightarrow{\mathrm{w}}\delta_0,\quad 其中\ S_n=\sum_{k=1}^{n}X_k.$$

**证** 令 $\mu_r^{(k)}=\boldsymbol{E}(|X_k|^r)$, $f_k(t)$ 为 $X_k$ 的特征函数, 则由第四章定理 3.4 有

$$f_k\left(\frac{t}{n}\right)=f_k(0)+f_k'(0)\frac{t}{n}+\theta_{n,k}2^{1-\delta}\mu_{1+\delta}^{(k)}\left(\frac{t}{n}\right)^{1+\delta}\frac{1}{1+\delta},$$

其中 $|\theta_{n,k}|\leqslant1$. 令

$$\omega_{n,k}(t)=\theta_{n,k}2^{1-\delta}\mu_{1+\delta}^{(k)}\left(\frac{t}{n}\right)^{1+\delta}\frac{1}{1+\delta},$$

则

$$f_k\left(\frac{t}{n}\right)=1+\omega_{n,k}(t).$$

但是

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\left\{\frac{\mu_{1+\delta}^{(k)}}{n^{1+\delta}}\right\}\leqslant\lim_{n\to\infty}\frac{1}{n^{1+\delta}}\sum_{k=1}^{n}\mu_{1+\delta}^{(k)}=0,$$

所以

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}|\omega_{n,k}(t)|=0\quad(\text{在 }|t|\leqslant T\text{上一致成立}).$$

因此, 当 $n$ 充分大时,

$$\begin{aligned}\log f_k\left(\frac{t}{n}\right)&=\log(1+\omega_{n,k}(t))\\&=\omega_{n,k}(t)+o(\omega_{n,k}(t))\quad(n\to\infty).\end{aligned}$$

所以

$$\begin{aligned}\boldsymbol{E}(\exp(itS_n/n))&=\prod_{i=1}^{n}\boldsymbol{E}(\exp(itX_k/n))\\&=\prod_{k=1}^{n}f_k\left(\frac{t}{n}\right)=\exp\left(\log\prod_{k=1}^{n}f_k\left(\frac{t}{n}\right)\right)\\&=\exp\left[\sum_{k=1}^{n}(\omega_{n,k}(t)+o(\omega_{n,k}(t)))\right].\end{aligned}$$

但是, 由定理假设知 (注意 $|\theta_{n,k}|\leqslant 1$)

$$\lim_{n\to\infty}\sum_{k=1}^{n}\omega_{n,k}(t)=\lim_{n\to\infty}\sum_{k=1}^{n}\theta_{n,k}\frac{2^{1-\delta}}{1+\delta}\,t^{1+\delta}\,\frac{\mu_{1+\delta}^{(k)}}{n^{1+\delta}}=0,$$

所以

$$\boldsymbol{E}(\exp(itS_n/n))\to 1.$$

定理证毕.

**定理 3.3**　设 $\{X_n,n\geqslant 1\}$ 是相互独立的随机变量序列, 对应的分布函数为 $\{F_n(x),n\geqslant 1\}$, 满足 $\boldsymbol{E}(X_n)\equiv 0$. 若

(1) $\lim\limits_{n\to\infty}\sum\limits_{k=1}^{n}\int_{|x|\geqslant n}F_k(dx)=0$;

(2) $\lim\limits_{n\to\infty}\frac{1}{n}\sum\limits_{k=1}^{n}\int_{|x|<n}xF_k(dx)=0$;

(3) $\lim\limits_{n\to\infty}\frac{1}{n^2}\sum\limits_{k=1}^{n}\int_{|x|<n}x^2F_k(dx)=0$,

则

$$P\circ\left(\frac{S_n}{n}\right)^{-1}\xrightarrow{\text{w}}\delta_0\quad\left(\text{其中 }S_n=\sum_{k=1}^{n}X_k\right).$$

**证**　作 R.V. $X_{n,k}$ 如下:

$$X_{n,k}=\begin{cases}X_k, & \text{当 }|X_k|<n,\\0, & \text{当 }|X_k|\geqslant n.\end{cases}$$

令

$$S_{n,n} = \sum_{k=1}^{n} X_{n,k} \quad (n \geqslant 1),$$

由于

$$\begin{aligned} P\left(\frac{S_n}{n} \neq \frac{S_{n,n}}{n}\right) &\leqslant \sum_{k=1}^{n} P(X_{n,k} \neq X_k) \\ &= \sum_{k=1}^{n} \int_{|x| \geqslant n} F_k(dx), \end{aligned}$$

所以, 由条件 (1) 和命题 3.1 知分布列 $\left\{P \circ \left(\frac{S_n}{n}\right)^{-1}, n \geqslant 1\right\}$ 与分布列 $\left\{P \circ \left(\frac{S_{n,n}}{n}\right)^{-1}, n \geqslant 1\right\}$ 等价. 因此, 由附注 3.2 得知: 为证定理 3.3, 只需证明

$$P \circ \left(\frac{S_{n,n}}{n}\right)^{-1} \xrightarrow{\mathrm{w}} \delta_0.$$

但是, 由条件 (2),

$$\begin{aligned} \frac{1}{n}\boldsymbol{E}(S_{n,n}/n) &= \frac{1}{n}\sum_{k=1}^{\infty} \boldsymbol{E}(X_{n,k}) \\ &= \frac{1}{n}\sum_{k=1}^{n} \int_{|x|<n} x F_k(dx) \to 0, \end{aligned}$$

因此, 为证定理 3.3, 又只需证明

$$\boldsymbol{E}\left[\exp\left(it\left(\frac{S_{n,n}}{n} - \boldsymbol{E}\left(\frac{S_{n,n}}{n}\right)\right)\right)\right] \to 1.$$

但是, 由条件 (3), 有

$$\begin{aligned} &\frac{1}{n^2}\sum_{k=1}^{n} \boldsymbol{E}\left((X_{n,k} - \boldsymbol{E}(X_{n,k}))^2\right) \\ \leqslant\ &\frac{1}{n^2}\sum_{k=1}^{n} \boldsymbol{E}\left(X_{n,k}^2\right) \\ \leqslant\ &\frac{1}{n^2}\sum_{k=1}^{n} \int_{|x|<n} x^2 F_k(dx) \to 0 \quad (n \to \infty), \end{aligned}$$

所以, 由定理 3.2 即得

$$\boldsymbol{E}\left[\exp\left(it\left(\frac{S_{n,n}}{n} - \boldsymbol{E}\left(\frac{S_{n,n}}{n}\right)\right)\right)\right] \to 1.$$

定理得证.

**推论 3.1** 设随机变量序列 $\{X_n, n \geqslant 1\}$ 相互独立, 其对应的分布函数列为 $\{F_n(x), n \geqslant 1\}$, $S_n = \sum\limits_{k=1}^{n} X_k$, $\boldsymbol{E}(X_n) = p_n$ 存在 $(n \geqslant 1)$. 若

(1)′ $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \int_{|x-p_k|\geqslant n} F_k(dx) = 0$;

(2)′ $\lim\limits_{n\to\infty} \dfrac{1}{n} \sum\limits_{k=1}^{n} \int_{|x-p_k|<n} (x-p_k) F_k(dx) = 0$;

(3)′ $\lim\limits_{n\to\infty} \dfrac{1}{n^2} \sum\limits_{k=1}^{n} \int_{|x-p_k|<n} (x-p_k)^2 F_k(dx) = 0$,

则

$$P \circ \left(\frac{S_n - \boldsymbol{E}(S_n)}{n}\right)^{-1} \stackrel{\mathrm{w}}{\Longrightarrow} \delta_0.$$

## §4 独立随机变量列的中心极限定理

**定理 4.1 (Liapunov 中心极限定理)** 设 $\{X_n, n \geqslant 1\}$ 为相互独立的随机变量序列, $\boldsymbol{E}(X_n) = 0$ $(\forall n \geqslant 1)$, 若对某个 $\delta > 0$, 有

$$\frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} \boldsymbol{E}\left(|X_k|^{2+\delta}\right) \to 0 \quad (n \to \infty), \tag{4.1}$$

则

$$P \circ \left(\frac{S_n}{s_n}\right)^{-1} \stackrel{\mathrm{w}}{\Longrightarrow} N(0,1),$$

其中 $S_n = \sum\limits_{k=1}^{n} X_k$, $s_n^2 = \mathrm{var}(S_n)$, $\sigma_k^2 = \mathrm{var}(X_k)$.

**证** 记 $\mu_{2+\delta}^{(k)} = \boldsymbol{E}\left(|X_k|^{2+\delta}\right)$, $f_k(t) = \boldsymbol{E}\left(e^{itX_k}\right)$.

(1) 先设存在 $0 < \delta \leqslant 1$ 使 (4.1) 成立.

由特征函数的 Taylor 展开有

$$\begin{aligned} f_k\left(\frac{t}{s_n}\right) &= 1 + f_k'(0)\frac{t}{s_n} + \frac{f_k''(0)}{2}\left(\frac{t}{s_n}\right)^2 \\ &\quad + \theta_{n,k} 2^{1-\delta} \mu_{2+\delta}^{(k)} \frac{1}{(1+\delta)(2+\delta)} \left(\frac{t}{s_n}\right)^{2+\delta} \\ &= 1 - \frac{\sigma_k^2}{2}\left(\frac{t}{s_n}\right)^2 + \theta_{n,k} 2^{1-\delta} \mu_{2+\delta}^{(k)} \frac{1}{(1+\delta)(2+\delta)} \left(\frac{t}{s_n}\right)^{2+\delta}, \quad (|\theta_{n,k}| \leqslant 1). \end{aligned}$$

令

$$\omega_{n,k}(t) = -\frac{\sigma_k^2}{2}\left(\frac{t}{s_n}\right)^2 + \theta_{n,k} 2^{1-\delta} \mu_{2+\delta}^{(k)} \frac{1}{(1+\delta)(2+\delta)} \left(\frac{t}{s_n}\right)^{2+\delta},$$

则由 (4.1) 得

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\left(\frac{\mu_{2+\delta}^{(k)}}{s_n^{2+\delta}}\right)\leqslant \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\mu_{2+\delta}^{(k)}}{s_n^{2+\delta}}=0. \tag{4.2}$$

由 Hölder 不等式有:

$$\sigma_k^{2+\delta}=\boldsymbol{E}(|X_k|^2)^{\frac{2+\delta}{2}}\leqslant \boldsymbol{E}\left(|X_k|^{2+\delta}\right)=\mu_{2+\delta}^{(k)},$$

所以

$$\frac{\sigma_k^2}{s_n^2}\leqslant\left(\frac{\sigma_k^{2+\delta}}{s_n^{2+\delta}}\right)^{\frac{2}{2+\delta}}\leqslant\left(\frac{\mu_{2+\delta}^{(k)}}{s_n^{2+\delta}}\right)^{\frac{2}{2+\delta}} \tag{4.3}$$

由 (4.2)、(4.3) 得:

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\frac{\sigma_k^2}{s_n^2}=0. \tag{4.4}$$

由 (4.3)、(4.4) 知:

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}|\omega_{n,k}(t)|=0\quad(\text{对 } |t|<T \text{ 一致成立}). \tag{4.5}$$

因此, 当 $n$ 充分大时, 对 $k=1,2,\cdots,n$,

$$\log\ f_k\left(\frac{t}{s_n}\right)=\log(1+\omega_{n,k}(t))$$

存在, 而且

$$\begin{aligned}\log f_k\left(\frac{t}{s_n}\right)&=\log(1+\omega_{n,k}(t))\\&=\omega_{n,k}(t)+o(\omega_{n,k}(t)),\quad(n\to\infty).\end{aligned}$$

所以

$$\begin{aligned}\boldsymbol{E}\left(\exp\left(\frac{itS_n}{s_n}\right)\right)&=\prod_{k=1}^{n}f_k\left(\frac{t}{s_n}\right)\\&=\exp\left[\sum_{k=1}^{n}(\omega_{n,k}(t)+o(\omega_{n,k}(t)))\right].\end{aligned} \tag{4.6}$$

但是由 (4.1) 知:

$$\begin{aligned}\sum_{k=1}^{n}\omega_{n,k}(t)&=\sum_{k=1}^{n}\left[-\frac{\sigma_k^2}{2}\left(\frac{t}{s_n}\right)^2+\frac{2^{1-\delta}\theta_{n,k}}{(1+\delta)(2+\delta)}t^{2+\delta}\frac{\mu_{2+\delta}^{(k)}}{s_n^{2+\delta}}\right]\\&=-\frac{1}{2}t^2+\frac{2^{1-\delta}}{(1+\delta)(2+\delta)}\sum_{k=1}^{n}\theta_{n,k}\frac{\mu_{2+\delta}^{(k)}}{s_n^{2+\delta}}t^{2+\delta}\\&\to-\frac{1}{2}t^2\quad(\text{当 } n\to\infty).\end{aligned} \tag{4.7}$$

由 (4.5)、(4.6)、(4.7) 得：

$$\boldsymbol{E}\left(\exp\left(\frac{itS_n}{s_n}\right)\right) \to e^{-\frac{1}{2}t^2},$$

即 $P \circ \left(\frac{S_n}{s_n}\right)^{-1} \xrightarrow{\text{w}} N(0,1)$.

(2) 再设存在 $\delta > 1$ 使 (4.1) 成立.

首先注意: 若 R.V. $Y$ 满足 $\boldsymbol{E}\left(|Y|^{2+\delta}\right) < \infty$, 则由 Hölder 不等式有:

$$\begin{aligned}\boldsymbol{E}\left(|Y|^3\right) &= \boldsymbol{E}\left(|Y|^{\frac{2+\delta}{\delta}} \bullet |Y|^{\frac{2\delta-2}{\delta}}\right) \\ &\leqslant \left[\boldsymbol{E}\left(|Y|^{2+\delta}\right)\right]^{\frac{1}{\delta}} \left[\boldsymbol{E}\left(|Y|^2\right)\right]^{\frac{\delta-1}{\delta}}.\end{aligned} \tag{4.8}$$

今作 R.V. $Y$ 具有分布函数 $\frac{1}{n}\sum\limits_{k=1}^{n} F_k(x)$, 其中 $F_k(x)$ 是 $X_k$ 的分布函数, 则对任何 $r$, 有:

$$\begin{aligned}\boldsymbol{E}\left(|Y|^r\right) &= \frac{1}{n}\sum_{k=1}^{n}\int_{\mathbf{R}} |x|^r F_k(dx) \\ &= \frac{1}{n}\sum_{k=1}^{n} \boldsymbol{E}\left(|X_k|^r\right).\end{aligned} \tag{4.9}$$

所以, 由 (4.8) 和 (4.9) 得

$$\begin{aligned}&\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^3\right) = \boldsymbol{E}(|Y|^3) \\ &\leqslant \left[\boldsymbol{E}\left(|Y|^{2+\delta}\right)\right]^{\frac{1}{\delta}} \left[\boldsymbol{E}(|Y|^2)\right]^{\frac{\delta-1}{\delta}} \\ &\leqslant \left[\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^{2+\delta}\right)\right]^{\frac{1}{\delta}} \left[\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^2\right)\right]^{\frac{\delta-1}{\delta}},\end{aligned}$$

所以

$$\begin{aligned}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^3\right) &\leqslant \left[\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^{2+\delta}\right)\right]^{\frac{1}{2}} \left[\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^2\right)\right]^{\frac{\delta-1}{\delta}} \\ &= \left[\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^{2+\delta}\right)\right]^{\frac{1}{\delta}} s_n^{\frac{2(\delta-1)}{\delta}}.\end{aligned} \tag{4.10}$$

把 (4.10) 两边除以 $s^3$ 并注意现在第 (2) 段假定了 (4.1) 对 $\delta$ 成立, 故

$$\frac{1}{s_n^3}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^3\right) \leqslant \left[\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{n}\boldsymbol{E}\left(|X_k|^{2+\delta}\right)\right]^{\frac{1}{\delta}} \to 0. \tag{4.11}$$

(4.11) 说明存在 $\delta'=1$, 而当 (4.1) 中的 $\delta$ 取为 $\delta'$ 时成立. 由第 (1) 段的证明知

$$P\circ\left(\frac{S_n}{s_n}\right)^{-1}\xrightarrow{\text{w}}N(0,1).$$

定理证毕.

**定理 4.2 (Lindeberg)** 设 $\{X_n,n\geqslant 1\}$ 是相互独立的随机变量序列, $\text{var}(X_n)=\sigma_n^2<\infty$ 存在, $\boldsymbol{E}(X_n)=0, F_n(x)$ 是 $X_n$ 的 d.f., $(n=1,2,\cdots)$. 令 $s_n^2=\sum\limits_{k=1}^n\sigma_k^2$, $S_n=\sum\limits_{k=1}^n X_k$, 则 $P\circ\left(\frac{S_n}{s_n}\right)^{-1}\xrightarrow{\text{w}}N(0,1)$ 而且 $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant n}\frac{\sigma_k^2}{s_n^2}=0$ 的充分必要条件是:

任给 $\varepsilon>0$, 都有

$$g_n(\varepsilon)\overset{\text{def.}}{=\!=}\frac{1}{s_n^2}\sum_{k=1}^n\int_{|x|\geqslant\varepsilon s_n}x^2F_k(dx)\to 0,\quad (\text{当 } n\to\infty). \tag{4.12}$$

(一般称条件 (4.12) 为 Lindeberg 条件.)

**证**　充分性. 若对每个 $\varepsilon$, 都有 $g_n(\varepsilon)\to 0$, 则对每个 $k$, 存在 $n_k$, 使 $n_{k+1}>n_k, g_n\left(\frac{1}{k}\right)\leqslant\frac{1}{k^3}$ $(\forall n\geqslant n_k)$. 取

$$\varepsilon_n=\frac{1}{k}\quad(\text{当 } n_k\leqslant n<n_{k+1}),$$

则有

$$\lim_{n\to\infty}\frac{1}{\varepsilon_n^2}g_n(\varepsilon_n)=0,\quad \lim_{n\to\infty}\frac{1}{\varepsilon_n}g_n(\varepsilon_n)=0. \tag{4.13}$$

显然,

$$\begin{aligned}&\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\frac{\sigma_k^2}{s_n^2}\\ &=\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\frac{1}{s_n^2}\left[\int_{|x|\geqslant\varepsilon_n s_n}x^2F_k(dx)-\int_{|x|<\varepsilon_n s_n}x^2F_k(dx)\right]\\ &\leqslant\lim_{n\to\infty}\left[g_n(\varepsilon_n)+\varepsilon_n^2\right]=0.\end{aligned} \tag{4.14}$$

再证 $P\circ\left(\frac{S_n}{s_n}\right)^{-1}\xrightarrow{\text{w}}N(0,1)$. 为此, 令

$$X_{n,k}=\begin{cases}X_k, & \text{当 } |X_k|<\varepsilon_n s_n,\\ 0, & \text{反之},\end{cases}$$

$$S_{n,n}=\sum_{k=1}^n X_{n,k}\quad(n\geqslant 1).$$

由于

$$\begin{aligned}P\left(\frac{S_n}{s_n}\neq\frac{S_{n,n}}{s_n}\right)&\leqslant\sum_{k=1}^{n}P(X_{n,k}\neq X_k)\\&=\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon_n s_n}F_k(dx)\leqslant\frac{1}{\varepsilon_n^2 s_n^2}\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon_n s_n}x^2F_k(dx)\\&=\frac{1}{\varepsilon_n^2}\bullet g_n(\varepsilon_n)\to 0\quad(\text{当 } n\to\infty),\end{aligned}$$

此即 $\left\{P\circ\left(\frac{S_n}{s_n}\right)^{-1},n\geqslant 1\right\}$ 与 $\left\{P\circ\left(\frac{S_{n,n}}{s_n}\right)^{-1},n\geqslant 1\right\}$ 等价. 因此, 由命题 3.1, 欲证

$$P\circ\left(\frac{S_n}{s_n}\right)^{-1}\stackrel{\text{w}}{\Longrightarrow}N(0,1),$$

只需证明:

$$P\circ\left(\frac{S_{n,n}}{s_n}\right)^{-1}\stackrel{\text{w}}{\Longrightarrow}N(0,1).\tag{4.15}$$

但是, 由 $\boldsymbol{E}(X_k)\equiv 0$ 及 $X_{n,k}$ 之定义知:

$$\begin{aligned}|\boldsymbol{E}(X_{n,k})|&=\left|\int_{|x|<\varepsilon_n s_n}xF_k(dx)\right|=\left|\int_{|x|\geqslant\varepsilon_n s_n}xF_k(dx)\right|\\&\leqslant\frac{1}{\varepsilon_n s_n}\left|\int_{|x|\geqslant\varepsilon_n s_n}x^2F_k(dx)\right|,\end{aligned}\tag{4.16}$$

所以由 $S_{n,n}$ 的定义 (4.16) 和 (4.13) 知

$$\begin{aligned}\left|\frac{\boldsymbol{E}(S_{n,n})}{s_n}\right|&\leqslant\frac{1}{s_n}\sum_{k=1}^{n}|\boldsymbol{E}(X_{n,k})|\leqslant\frac{1}{\varepsilon_n s_n^2}\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon_n s_n}x^2F_k(dx)\\&=\frac{1}{\varepsilon_n}g_n(\varepsilon_n)\to 0\quad(\text{当 } n\to\infty).\end{aligned}\tag{4.17}$$

因此

$$\boldsymbol{E}\left(e^{it\boldsymbol{E}(S_{n,n})/s_n}\right)\to 1.$$

所以, 欲证 $P\circ\left(\frac{S_{n,n}}{s_n}\right)^{-1}\stackrel{\text{w}}{\Longrightarrow}N(0,1)$, 又只需证

$$\boldsymbol{E}\left(\exp[it(S_{n,n}-\boldsymbol{E}(S_{n,n}))/s_n]\right)\to e^{-\frac{1}{2}t^2}\quad(\text{当 } n\to\infty).\tag{4.18}$$

令

$$s_{n,n}^2=\operatorname{var}(S_{n,n})=\sum_{k=1}^{n}\operatorname{var}(X_{n,k}),\tag{4.19}$$

由 $\boldsymbol{E}(X_k)\equiv 0$ 得

$$
\begin{aligned}
&1-\frac{s_{n,n}^2}{s_n^2}\\
=&\frac{1}{s_n^2}\left[\sum_{k=1}^n\int_{\mathbf{R}}x^2F_k(dx)-\sum_{k=1}^n\left\{\int_{|x|<\varepsilon_n s_n}x^2F_k(dx)-\left(\int_{|x|<\varepsilon_n s_n}xF_k(dx)\right)^2\right\}\right]\\
=&\frac{1}{s_n^2}\left[\sum_{k=1}^n\int_{|x|\geqslant\varepsilon_n s_n}x^2F_k(dx)+\sum_{k=1}^n\left(\int_{|x|<\varepsilon_n s_n}xF_k(dx)\right)^2\right]\\
=&g_n(\varepsilon_n)+\frac{1}{s_n^2}\sum_{k=1}^n\left(\int_{|x|<\varepsilon_n s_n}xF_k(dx)\right)^2\\
\leqslant&g_n(\varepsilon_n)+\frac{1}{s_n^2}\cdot\frac{1}{\varepsilon_n^2 s_n^2}\sum_{k=1}^n\left(\int_{|x|<\varepsilon_n s_n}x^2F_k(dx)\right)^2\\
\leqslant&g_n(\varepsilon_n)+\frac{1}{\varepsilon_n^2}\left(\frac{1}{s_n^2}\sum_{k=1}^n\int_{|x|<\varepsilon_n s_n}x^2F_k(dx)\right)^2\\
=&g_n(\varepsilon_n)+\frac{1}{\varepsilon_n^2}g_n(\varepsilon_n)^2.
\end{aligned}
\tag{4.20}
$$

由 (4.20) 和 (4.13) 知

$$
\lim_{n\to\infty}\frac{s_{n,n}^2}{s_n^2}=1. \tag{4.21}
$$

而

$$
\begin{aligned}
&\frac{1}{s_{n,n}^3}\sum_{k=1}^n\boldsymbol{E}\left(|X_{n,k}-\boldsymbol{E}(X_{n,k})|^3\right)\\
=&\frac{1}{s_{n,n}^3}\sum_{k=1}^n\left(\int_{|x|<\varepsilon_n s_n}|x-\boldsymbol{E}(X_{n,k})|^3F_k(dx)+|\boldsymbol{E}(X_{n,k})|^3\int_{|x|\geqslant\varepsilon_n s_n}F_k(dx)\right)\\
\leqslant&\frac{2\varepsilon_n s_n}{s_{n,n}^3}\sum_{k=1}^n\int_{|x|<\varepsilon_n s_n}|x-\boldsymbol{E}(X_{n,k})|^2F_k(dx)+\frac{\varepsilon_n s_n^3}{s_{n,n}^3}g_n(\varepsilon_n)\\
\leqslant&\frac{2\varepsilon_n s_n}{s_{n,n}}\cdot\frac{1}{s_{n,n}^2}\sum_{k=1}^n\boldsymbol{E}\left(|X_{n,k}-\boldsymbol{E}(X_{n,k})|^2\right)+\frac{\varepsilon_n s_n^3}{s_{n,n}^3}g_n(\varepsilon_n)\\
=&\frac{2\varepsilon_n s_n}{s_{n,n}}+\frac{\varepsilon_n s_n^3}{s_{n,n}^3}g_n(\varepsilon_n),
\end{aligned}
\tag{4.22}
$$

所以由 (4.21)、(4.22) 和 (4.13) 知:

$$
\lim_{n\to\infty}\frac{1}{s_{n,n}^3}\sum_{k=1}^n\boldsymbol{E}(|X_{n,k}-\boldsymbol{E}(X_{n,k})|^3)=0. \tag{4.23}
$$

由 (4.23) 及定理 4.1 知

$$\boldsymbol{E}\left(\exp[it(S_{n,n}-\boldsymbol{E}(S_{n,n}))/s_{n,n}]\right)\to e^{-\frac{1}{2}t^2}. \tag{4.24}$$

再用 $\lim\limits_{n\to\infty}\dfrac{s_{n,n}}{s_n}=1$ 及第四章推论 2.2, 由 (4.24) 可得 (4.18). 定理的充分性部分证毕.

必要性. 设

$$P\circ\left(\frac{S_n}{s_n}\right)^{-1}\xrightarrow{\mathrm{w}}N(0,1),\quad 且\quad \lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\frac{\sigma_k^2}{s_n^2}=0.$$

令 $f_k(t)$ 为 $X_k$ 的 c.f.. 则由

$$\begin{aligned}f_k\left(\frac{t}{s_n}\right)&=1+f_k'(0)\frac{t}{s_n}+\theta f''(0)\frac{1}{2}\left(\frac{t}{s_n}\right)^2\\&=1-\theta\frac{\sigma_k^2}{2}\left(\frac{t}{s_n}\right)^2\quad(|\theta|\leqslant 1)\end{aligned}$$

可推出

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\left|f_k\left(\frac{t}{s_n}\right)-1\right|\leqslant\lim_{n\to\infty}\max_{1\leqslant k\leqslant n}\left|\frac{\sigma_k^2}{s_n^2}\frac{t^2}{2}\right|=0. \tag{4.25}$$

因此, 当 $n$ 充分大, $1\leqslant k\leqslant n$ 时, $\log f_k\left(\dfrac{t}{s_n}\right)$ 存在. 由假设知

$$\lim_{n\to\infty}\prod_{k=1}^{n}f_k\left(\frac{t}{s_n}\right)=e^{-\frac{1}{2}t^2}. \tag{4.26}$$

把 (4.26) 换成对数形式:

$$\lim_{n\to\infty}\sum_{k=1}^{n}\log f_k\left(\frac{t}{s_n}\right)=-\frac{1}{2}t^2,$$

即

$$\lim_{n\to\infty}\sum_{k=1}^{n}\left\{\left(f_k\left(\frac{t}{s_n}\right)-1\right)+M_{n,k}(t)\left(f_k\left(\frac{t}{s_n}\right)-1\right)^2\right\}=-\frac{1}{2}t^2,$$

其中

$$\max_{1\leqslant k\leqslant n}|M_{n,k}(t)|\leqslant K(t)<\infty\quad(\forall n\geqslant 1).$$

而

$$\begin{aligned}\sum_{k=1}^{n}\left|f_k\left(\frac{t}{s_n}\right)-1\right|^2&\leqslant\max_{1\leqslant k\leqslant n}\left|f_k\left(\frac{t}{s_n}\right)-1\right|\sum_{k=1}^{n}\frac{\sigma_k^2}{2}\left(\frac{t}{s_n}\right)^2\\&=\frac{t^2}{2}\max_{1\leqslant k\leqslant n}\left|f_k\left(\frac{t}{s_n}\right)-1\right|\to 0\ (当\ n\to\infty),\end{aligned}$$

所以

$$\lim_{n\to\infty}\sum_{k=1}^{n}\left(f_k\left(\frac{t}{s_n}\right)-1\right)=-\frac{1}{2}t^2. \tag{4.27}$$

对 (4.27) 取实部即得

$$\lim_{n\to\infty}\sum_{k=1}^{n}\int_{\mathbf{R}}\left(\cos\frac{tx}{s_n}-1\right)F_k(dx)=-\frac{1}{2}t^2,$$

即

$$\sum_{k=1}^{n}\int_{\mathbf{R}}\left(\cos\frac{tx}{s_n}-1\right)F_k(dx)=-\frac{1}{2}t^2+o(1)\quad(\text{当 } n\to\infty). \tag{4.28}$$

但是

$$\begin{aligned}
&\sum_{k=1}^{n}\int_{|x|<\varepsilon s_n}\left(1-\cos\frac{tx}{s_n}\right)F_k(dx)\\
\leqslant&\sum_{k=1}^{n}\int_{|x|<\varepsilon s_n}\frac{1}{2}\left(\frac{tx}{s_n}\right)^2F_k(dx)\\
\leqslant&\frac{t^2}{2s_n^2}\sum_{k=1}^{n}\left(\int_{\mathbf{R}}x^2F_k(dx)-\int_{|x|>\varepsilon s_n}x^2F_k(dx)\right)\\
=&\frac{t^2}{2s_n^2}\sum_{k=1}^{n}\left(\sigma_k^2-\int_{|x|>\varepsilon s_n}x^2F_k(dx)\right)\\
=&\frac{t^2}{2}(1-g_n(\varepsilon)),
\end{aligned} \tag{4.29}$$

又

$$\begin{aligned}
\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon s_n}\left(1-\cos\frac{tx}{s_n}\right)F_k(dx)&\leqslant 2\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon s_n}F_k(dx)\\
&\leqslant\frac{2}{\varepsilon^2 s_n^2}\sum_{k=1}^{n}\int_{|x|\geqslant\varepsilon s_n}x^2F_k(dx)\\
&=\frac{2}{\varepsilon^2}g_n(\varepsilon)\leqslant\frac{2}{\varepsilon^2}.
\end{aligned} \tag{4.30}$$

把 (4.29)、(4.30) 代入 (4.28) 得

$$o(1)+\frac{t^2}{2}\leqslant\frac{t^2}{2}(1-g_n(\varepsilon))+\frac{2}{\varepsilon^2}\quad(\text{当 } n\to\infty),$$

即

$$o\leqslant g_n(\varepsilon)\leqslant\frac{2}{t^2}\left(\frac{2}{\varepsilon^2}+o(1)\right)\quad(\text{当 } n\to\infty).$$

先令 $n \to \infty$, 再令 $t \to \infty$ 即得

$$\lim_{n\to\infty} g_n(\varepsilon) = 0 \quad (\forall \varepsilon > 0).$$

定理证毕.

**推论 4.1** 若把定理 4.1 中的条件 $\boldsymbol{E}(X_n) = 0$ $(\forall n \geqslant 1)$ 改为 $\boldsymbol{E}(X_n) = p_n$ 存在 $(\forall n \geqslant 1)$, 把条件 (4.1) 改为

$$\frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} \boldsymbol{E}\left(|X_k - \boldsymbol{E}(X_k)|^{2+\delta}\right) \to 0 \quad (n \to \infty), \tag{4.1$'$}$$

而其他条件不变, 则有

$$P \circ \left(\frac{S_n - \boldsymbol{E}(S_n)}{s_n}\right)^{-1} \xrightarrow{\text{w}} N(0,1).$$

**推论 4.2** 若把定理 4.2 中的条件 $\boldsymbol{E}(X_n) = 0$ $(\forall n \geqslant 1)$ 改为 $\boldsymbol{E}(|X_n|) = p_n$ 存在 $(\forall n \geqslant 1)$, 把条件 (4.12) 改为

$$g_n(\varepsilon)^* \overset{\text{def.}}{=\!=} \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{|x-p_k| \geqslant \varepsilon s_n} (x - p_k)^2 F_k(dx) \to 0 \quad (n \to \infty), \tag{4.12$^*$}$$

而其他条件不变, 则有

$$《 P \circ \left(\frac{S_n - \boldsymbol{E}(S_n)}{s_n}\right)^{-1} \xrightarrow{\text{w}} N(0,1), \quad \lim_{n\to\infty} \max_{1 \leqslant k \leqslant n} \frac{\sigma_k^2}{s_n^2} = 0 》$$

$\Leftrightarrow$ (4.12)$^*$ 成立.

以后对这种小技巧不再解释了. 在概率极限理论中, 假定 $\boldsymbol{E}(X_k) = 0$ 与 $\boldsymbol{E}(X_k)$ 存在, 无本质差异.

## §5 强大数定律和随机级数的收敛性

在 §1 和 §3 中所讨论的大数定律是随机变量序列 $\{X_k, k \geqslant 1\}$ 前 $n$ 项的算术平均的 "正则化" $\frac{1}{n}\sum\limits_{k=1}^{n}(X_k - p_k)$ 的依概率收敛的问题, 这类大数定律也称为弱大数定律. 在 §5 中, 我们将要研究的强大数定律, 即 $\frac{1}{n}\sum\limits_{k=1}^{n}(X_k - p_k)$ a.s. 收敛的问题. 本节中, 完备概率空间 $(\Omega, \mathscr{F}, P)$ 是给定的.

**引理 5.1 (Kolmogorov 不等式)** 设随机变量 $X_1, \cdots, X_n$ 相互独立, $\boldsymbol{E}(X_k) = p_k, \operatorname{var}(X_k) = \sigma_k^2$ 都存在 $(k = 1, \cdots, n)$, 则对任何 $\varepsilon > 0$ 都有

$$P\left(\max_{1 \leqslant k \leqslant n} \left|\sum_{i=1}^{k}(X_i - p_i)\right| \geqslant \varepsilon\right) \leqslant \frac{\sum\limits_{k=1}^{n} \sigma_k^2}{\varepsilon^2}. \tag{5.1}$$

**证** 令

$$S_k = \sum_{i=1}^{k}(X_i - p_i),\ B = \left\{\max_{1\leqslant k\leqslant n} |S_k| \geqslant \varepsilon\right\},$$

$$B_k = \left(\bigcap_{i=1}^{k-1}\{|S_i| < \varepsilon\}\right)\bigcap\{|S_k| \geqslant \varepsilon\}, \quad k = 1, \cdots, n$$

$\left(\bigcap\limits_{i=1}^{0} C_i\right.$ 定义为全空间, 即 $\left.B_1 = \{|S_1| \geqslant \varepsilon\}\right)$, 则由 $\{B_1, \cdots, B_n\}$ 两两不交, $\bigcup\limits_{k=1}^{n} B_k = B$ 及 $\{X_1, \cdots, X_n\}$ 相互独立可得

$$\begin{aligned}
\int_{B_k} S_n^2 dP &= \int_{B_k} \left[S_k^2 + 2S_k \sum_{i=k+1}^{n}(X_i - p_i) + \left(\sum_{i=k+1}^{n}(X_i - p_i)\right)^2\right] dP \\
&= \int_{B_k} S_k^2 dP + \int_{B_k} \left(\sum_{i=k+1}^{n}(X_i - p_i)\right)^2 dP \\
&\geqslant \int_{B_k} S_k^2 dP.
\end{aligned} \tag{5.2}$$

所以由 $\{X_1, \cdots, X_n\}$ 相互独立及 (5.2) 得:

$$\begin{aligned}
\sum_{k=1}^{n}\sigma_k^2 &= \mathrm{var}\left(\sum_{i=1}^{n} X_i\right) = \int_{\Omega} S_n^2 dP \\
&\geqslant \int_{B} S_n^2 dP = \sum_{k=1}^{n}\int_{B_k} S_n^2 dP \\
&\geqslant \sum_{k=1}^{n}\int_{B_k} S_k^2 dP \geqslant \sum_{k=1}^{n}\varepsilon^2 P(B_k) = \varepsilon^2 P(B),
\end{aligned}$$

此即 (5.1) 成立. 引理证毕.

**定理 5.1** 设随机变量序列 $\{X_n, n \geqslant 1\}$ 相互独立, $\boldsymbol{E}(X_n) = p_n, \mathrm{var}(X_n) = \sigma_n^2$ 都存在, $(n = 1, 2, \cdots)$. 则随机级数 $\sum\limits_{n=1}^{\infty} X_n$ $L^2-$ 收敛 (即存在平方可积的随机变量 $X$, 使 $\boldsymbol{E}\left(\left|\left(\sum\limits_{k=1}^{n} X_k\right) - X\right|^2\right) \to 0$) 的充分必要条件是:

$$\sum_{n=1}^{\infty} p_n \quad \text{与} \quad \sum_{n=1}^{\infty}\sigma_n^2 \quad \text{都收敛.}$$

此外, 若 $\sum\limits_{k=1}^{n} X_k \xrightarrow{L^2} X$, 则 $\sum\limits_{n=1}^{\infty} p_n = \boldsymbol{E}(X), \sum\limits_{n=1}^{\infty}\sigma_n^2 = \mathrm{var}(X)$.

**证** 必要性. 设

$$\sum_{k=1}^{n} X_k \xrightarrow{L^2} X,$$

则

$$\lim_{n\to\infty} \boldsymbol{E}\left(\left|\left(\sum_{k=1}^{n} X_k\right) - X\right|^2\right) = 0. \tag{5.3}$$

由

$$\begin{aligned}\left|\boldsymbol{E}(X) - \sum_{k=1}^{n} p_k\right| &\leqslant \boldsymbol{E}\left(\left|X - \sum_{k=1}^{n} X_k\right|\right)\\ &\leqslant \left[\boldsymbol{E}\left(\left|X - \sum_{k=1}^{n} X_k\right|^2\right)\right]^{\frac{1}{2}}\end{aligned}$$

立得

$$\sum_{n=1}^{\infty} p_n = \boldsymbol{E}(X) \text{ 收敛.}$$

又因为由 $\{X_n, n \geqslant 1\}$ 的相互独立性有

$$\begin{aligned}&\left|\operatorname{var}(X) - \sum_{k=1}^{n} \operatorname{var}(X_k)\right| = \left|\operatorname{var}(X) - \operatorname{var}\left(\sum_{k=1}^{n} X_k\right)\right|\\ &= \left|\boldsymbol{E}(X^2) - \boldsymbol{E}(X)^2 - \boldsymbol{E}\left(\left(\sum_{k=1}^{n} X_k\right)^2\right) + \left(\sum_{k=1}^{n} p_k\right)^2\right|\\ &\leqslant \boldsymbol{E}\left(\left|X - \sum_{k=1}^{n} X_k\right|^2\right) + 2\boldsymbol{E}\left(\left|\left(\sum_{k=1}^{n} X_k\right)^2 - X\sum_{k=1}^{n} X_k\right|\right)\\ &\quad + \left|\left(\sum_{k=1}^{n} p_k\right)^2 - \boldsymbol{E}(X)^2\right|.\end{aligned} \tag{5.4}$$

但是由 Hölder 不等式和 Minkowski 不等式有

$$\begin{aligned}&\boldsymbol{E}\left(\left|\left(\sum_{k=1}^{n} X_k\right)^2 - X\sum_{k=1}^{n} X_k\right|\right)\\ &\leqslant \boldsymbol{E}\left(\left|\sum_{k=1}^{n} X_k - X\right|^2\right)^{\frac{1}{2}} \boldsymbol{E}\left(\left|\sum_{k=1}^{n} X_k\right|^2\right)^{\frac{1}{2}}\\ &\leqslant \boldsymbol{E}\left(\left|\sum_{k=1}^{n} X_k - X\right|^2\right)^{\frac{1}{2}} \left[\boldsymbol{E}\left(\left|\sum_{k=1}^{n} X_k - X\right|^2\right)^{\frac{1}{2}} + \boldsymbol{E}(|X|^2)^{\frac{1}{2}}\right].\end{aligned} \tag{5.5}$$

以 (5.5) 代入 (5.4) 并注意 $\boldsymbol{E}(|X|^2)=\beta^2<\infty$, $\sum\limits_{k=1}^{\infty}p_k=\boldsymbol{E}(X)$ 收敛及 (5.3) 可得

$$\sum_{n=1}^{\infty}\sigma_n^2=\operatorname{var}(X)\quad 收敛.$$

充分性. 事实上, 由

$$\sum_{n=1}^{\infty}\sigma_n^2<\infty\quad 及\quad \boldsymbol{E}\left(\left|\sum_{k=1}^{n}(X_k-p_k)\right|^2\right)\leqslant\sum_{k=1}^{n}\sigma_k^2$$

可知

$$\left\{\sum_{k=1}^{n}(X_k-p_k),n\geqslant 1\right\}$$

是一个基本 $L^2-$ 收敛序列, 从而它是 $L^2$ 收敛序列. 又因为 $\sum\limits_{n=1}^{\infty}p_n$ 收敛, 所以 $\left\{\sum\limits_{k=1}^{n}X_k,n\geqslant 1\right\}$ $L^2-$ 收敛.

**定理 5.2** 设随机变量序列 $\{X_n,n\geqslant 1\}$ 相互独立, $\boldsymbol{E}(X_n)=p_n,\operatorname{var}(X_n)=\sigma_n^2$ 存在, $\sum\limits_{n=1}^{\infty}p_n$ 和 $\sum\limits_{n=1}^{\infty}\sigma_n^2$ 收敛, 则

(1) $\sum\limits_{n=1}^{\infty}X_n$ $L^2-$ 收敛, 且 $\boldsymbol{E}\left(\sum\limits_{n=1}^{\infty}X_n\right)=\sum\limits_{n=1}^{\infty}p_n,\operatorname{var}\left(\sum\limits_{n=1}^{\infty}X_n\right)=\sum\limits_{n=1}^{\infty}\sigma_n^2$;

(2) $\sum\limits_{n=1}^{\infty}X_n$ a.s. 收敛;

(3) $P\left(\sup\limits_{1\leqslant n<\infty}\left|\sum\limits_{k=1}^{n}(X_k-p_k)\right|\geqslant\varepsilon\right)\leqslant\sum\limits_{n=1}^{\infty}\sigma_n^2/\varepsilon^2$ ($\forall\varepsilon>0$ 成立).

**证** (1) 可由定理 5.1 立即可得.

(2) 不失普遍性可令 $p_n=0$ $(n=1,2,\cdots)$. 再令

$$S_0\equiv 0,\quad S_n=\sum_{k=1}^{n}X_k,$$
$$a_m=\sup\{|S_{m+k}-S_m|:k=1,2,\cdots\},$$
$$a=\inf\{a_m:m=0,1,2,\cdots\},$$

则对任何 $\omega_0\in\Omega$, $\sum\limits_{n=1}^{\infty}X_n(\omega_0)$ 收敛的充分必要条件是 $a(\omega_0)=0$. 因此, 欲证

$$\sum_{n=1}^{\infty}X_n\quad \text{a.s. 收敛},$$

只需证明对任何 $\varepsilon>0$, 有

$$P(a\geqslant\varepsilon)=0. \tag{5.6}$$

事实上,

$$\begin{aligned}
&P(a \geqslant \varepsilon) \leqslant P(a_m \geqslant \varepsilon) \\
&= P(\sup\{|S_{m+k} - S_m| : k = 1, 2, \cdots\} \geqslant \varepsilon) \\
&\leqslant P\left(\lim_{n\to\infty}\left(\max_{1\leqslant k\leqslant n}|S_{m+k} - S_m| \geqslant \varepsilon - \frac{\varepsilon}{r}\right)\right) \\
&= \lim_{n\to\infty} P\left(\max_{1\leqslant k\leqslant n}|S_{m+k} - S_m| \geqslant \left(\varepsilon - \frac{\varepsilon}{r}\right)\right) \\
&\leqslant \lim_{n\to\infty}\frac{1}{\left(\varepsilon - \frac{\varepsilon}{r}\right)^2}\sum_{i=m+1}^{m+n}\sigma_i^2,
\end{aligned} \tag{5.7}$$

由于 $\sum\limits_{n=1}^{\infty}\sigma_n^2 < \infty$, 所以, 在 (5.7) 中令 $m \to \infty$ 即得 (5.6). 故 (2) 成立.

(3) 在 (5.7) 中取 $m = 0$ 并令 $r \to \infty$ 即得 (3). 定理证毕.

**定理 5.3** 设 $\{X_n, n \geqslant 1\}$ 为相互独立的随机变量序列, 且 $|X_n| < C$, a.s. $(\forall n \geqslant 1)$. 若 $\sum\limits_{n=1}^{\infty} X_n$ 在某一个 $P$ 正测集上收敛, 则

$$\sum_{n=1}^{\infty}\boldsymbol{E}(X_n) \text{ 和 } \sum_{n=1}^{\infty}\operatorname{var}(X_n) \text{ 皆收敛.}$$

**证** 先设 $\boldsymbol{E}(X_n) = 0$ $(n \geqslant 1)$. 令 $S_0 \equiv 0, S_n = \sum\limits_{k=1}^{n} X_k$ $(n \geqslant 1)$. 由定理的假设知: 存在 $B_1$ 使 $P(B_1) > 0$ 且 $\sum\limits_{n=1}^{\infty} X_n(\omega)$ 收敛, $|X_n(\omega)| < C$ $(\forall \omega \in B_1, n \geqslant 1)$. 再用 Егоров 定理可知: 存在一个 $B_2 \subset B_1, P(B_2) > 0$, 使 $\sum\limits_{n=1}^{\infty} X_n(\omega)$ 在 $\omega \in B_2$ 上一致收敛. 所以, 任给 $\varepsilon > 0$, 存在正整数 $N$, 使

$$|S_{N+m}^{(\omega)} - S_N(\omega)| < \varepsilon \quad (\forall \omega \in B_2, m \geqslant 0),$$

故

$$|S_{N+m}(\omega)| \leqslant |S_N(\omega)| + \varepsilon \leqslant Nc + \varepsilon \quad (\forall \omega \in B_2, m \geqslant 0). \tag{5.8}$$

又显然有

$$|S_n(\omega)| \leqslant Nc + \varepsilon \quad (\forall n \leqslant N, \omega \in B_2). \tag{5.9}$$

由 (5.8) 和 (5.9) 得知

$$|S_n(\omega)| \leqslant Nc + \varepsilon \quad (\forall n \geqslant 1, \omega \in B_2). \tag{5.10}$$

即

$$B_2 \subset \bigcap_{n=0}^{\infty}\{|S_n| \leqslant Nc + \varepsilon\} \xlongequal{\text{记作}} D.$$

令

$$D_n = \bigcap_{k=0}^{n} \{|S_k| \leqslant Nc + \varepsilon\} \quad (n \geqslant 0),$$

则 $D_n \downarrow D, P(D) > 0$. 再令

$$d = Nc + \varepsilon, \quad F_n = D_{n-1} - D_n, \quad \alpha_n = \int_{D_n} S_n^2 dP,$$

则由 $\{X_n, n \geqslant 1\}$ 相互独立及假设 $\boldsymbol{E}(X_n) = 0$ 得:

$$\begin{aligned}
\alpha_n - \alpha_{n-1} &= \int_{D_n} S_n^2 dP - \int_{D_{n-1}} S_{n-1}^2 dP \\
&= \int_{D_{n-1}} S_n^2 dP - \int_{F_n} S_n^2 dP - \int_{D_{n-1}} S_{n-1}^2 dP \\
&= \int_{D_{n-1}} (X_n^2 + 2X_n S_{n-1}) dP - \int_{F_n} S_n^2 dP \\
&= \int_{D_{n-1}} X_n^2 dP - \int_{F_n} S_n^2 dP \\
&= \int_{\Omega} \mathbf{1}_{D_{n-1}} X_n^2 dP - \int_{F_n} S_n^2 dP \\
&= \int_{\Omega} \mathbf{1}_{D_{n-1}} dP \cdot \int_{\Omega} X_n^2 dP - \int_{F_n} S_n^2 dP \\
&\geqslant P(D)\mathrm{var}(X_n) - P(F_n)(d+c)^2.
\end{aligned} \tag{5.11}$$

把 (5.11) 两边对 $n$ 从 1 到 $M$ 求和得

$$\alpha_M \geqslant \sum_{n=1}^{M} P(D)\mathrm{var}(X_n) - (d+c)^2. \tag{5.12}$$

由 $\alpha_M \leqslant d^2, (\forall M \geqslant 1), P(D) > 0$ 及 (5.12) 可知

$$\sum_{n=1}^{\infty} \mathrm{var}(X_n) < \infty.$$

现在取消 $\boldsymbol{E}(X_n) = 0$ 的假设. 作新的随机变量序列 $\{X_n^*, n \geqslant 1\}$, 使之与 $\{X_n, n \geqslant 1\}$ 相互独立, 且 $X_n^*$ 与 $X_n$ 之分布相同 (利用乘积空间技巧, 这样的随机变量序列是存在的). 令 $Y_n = X_n - X_n^*$ $(n \geqslant 1)$, 则 $\{Y_n, n \geqslant 1\}$ 相互独立, $\boldsymbol{E}(Y_n) = 0, |Y_n| < 2C$, a.s. $(\forall n \geqslant 1)$. 因此由前面的证明知 $\sum\limits_{n=1}^{\infty} \mathrm{var}(Y_n) < \infty$, 从而 $\sum\limits_{n=1}^{\infty} \mathrm{var}(X_n) = \dfrac{1}{2} \sum\limits_{n=1}^{\infty} \mathrm{var}(Y_n) < \infty$. 所以

$$\sum_{n=1}^{\infty} \mathrm{var}(X_n - \boldsymbol{E}(X_n)) < \infty, \quad \sum_{n=1}^{\infty} \boldsymbol{E}(X_n - \boldsymbol{E}(X_n)) = 0.$$

因此, 由定理 5.1 知

$$\sum_{n=1}^{\infty}(X_n - \boldsymbol{E}(X_n)) \quad \text{a.s.} \quad 收敛,$$

而 $\sum\limits_{n=1}^{\infty} X_n$ 在某个正测集上收敛, 所以 $\sum\limits_{n=1}^{\infty} \boldsymbol{E}(X_n)$ 收敛. 定理证毕.

下面我们要证明强大数定律, 为此, 先证明几个引理.

**引理 5.2 (Borel – Contelli 引理)** 设 $\{B_n, n \geqslant 1\}$ 是一列 $\mathscr{F}$ 中的可测集.

(1) 若 $\sum\limits_{n=1}^{\infty} P(B_n) < \infty$, 则 $P\left(\lim\limits_{n\to\infty} \sup B_n\right) = 0$;

(2) 若 $P\left(\lim\limits_{n\to\infty} \sup B_n\right) = 0$ 且 $\{B_n, n \geqslant 1\}$ 相互独立, 则 $\sum\limits_{n=1}^{\infty} P(B_n) < \infty$.

**证** (1) 设 $\sum\limits_{n=1}^{\infty} P(B_n) < \infty$, 则

$$P\left(\lim_{n\to\infty} \sup B_n\right) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k\right)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} B_k\right) \leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} P(B_k) = 0.$$

(2) 设 $P\left(\lim\limits_{n\to\infty} \sup B_n\right) = 0$ 且 $\{B_n, n \geqslant 1\}$ 相互独立, 则

$$\omega_0 \in \lim_{n\to\infty} \sup B_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} B_k$$

的充分必要条件是

$$\sum_{n=1}^{\infty} \mathbf{1}_{B_n}(\omega_0) \quad 发散.$$

所以, 由 $P\left(\lim\limits_{n\to\infty} \sup B_n\right) = 0$ 可以推出

$$\sum_{n=1}^{\infty} \mathbf{1}_{B_n} \quad \text{a.s.} \quad 收敛.$$

而 $|\mathbf{1}_{B_n}(\omega)| \leqslant 1 \quad (\forall \omega \in \Omega, n \geqslant 1)$, 所以由定理 5.3 知:

$$\sum_{n=1}^{\infty} P(B_n) = \sum_{n=1}^{\infty} \boldsymbol{E}(\mathbf{1}_{B_n}) \quad 收敛.$$

**引理 5.3** 设实数矩阵 $\{a_{n,k}, 1 \leqslant k \leqslant k_n, n \geqslant 1\}$ 满足

(1) $k_{n+1} \geqslant k_n \geqslant n \ (\forall n \geqslant 1)$;

(2) 对每个 $k \geqslant 1$, 都有 $\lim\limits_{n\to\infty} a_{n,k} = 0$;

(3) $\sum\limits_{k=1}^{k_n} |a_{n,k}| \leqslant c < \infty \ (\forall n \geqslant 1)$.

再令 $\{x_n, n \geqslant 1\}$ 是实数列, $x'_n = \sum\limits_{k=1}^{k_n} a_{n,k} x_k \ (n \geqslant 1)$, 则

1° $x_n \to 0 \Rightarrow x'_n \to 0$;

2° "$x_n \to x$ ($x$ 是实数) 和 $\sum\limits_{k=1}^{k_n} a_{n,k} \to 1$" $\Rightarrow x'_n \to x$. 特别地, 若

$$a_{n,k} = a_k, k_n = n,$$
$$b_n \overset{\text{def.}}{=\!=} \left(\sum_{k=1}^{n} a_k\right) \uparrow \infty \quad (\text{当 } n \uparrow \infty),$$

则

$$\text{“}x_n \to x \Rightarrow \frac{1}{b_n}\sum_{k=1}^{n} a_k x_k \to x.\text{”}$$

**证** 1° 设 $x_n \to 0$, 则对任何 $\varepsilon > 0$, 存在正整数 $N(\varepsilon)$, 当 $n \geqslant N(\varepsilon)$ 时有 $|x_n| \leqslant \dfrac{\varepsilon}{c}$. 所以

$$|x'_n| \leqslant \sum_{k=1}^{N(\varepsilon)} |a_{n,k} x_k| + \varepsilon. \tag{5.13}$$

由于 $\lim\limits_{n\to\infty} a_{n,k} = 0 \ (\forall k \geqslant 1)$, 所以在 (5.13) 中令 $n \to \infty$ 得

$$\lim_{n\to\infty} \sup |x'_n| \leqslant \varepsilon. \tag{5.14}$$

由 $\varepsilon > 0$ 可以取得任意小, 由 (5.14) 知 $x'_n \to 0$.

2° 若

$$\sum_{k=1}^{k_n} a_{n,k} \to 1, \quad x_n \to x \quad (\text{当 } n \to \infty),$$

则

$$x'_n = \left(\sum_{k=1}^{k_n} a_{n,k} x + \sum_{k=1}^{k_n} a_{n,k}(x_k - x)\right) \to x \quad (\text{当 } n \to \infty).$$

特别地, 若 $k_n = n, a_{n,k} = a_k, x_n \to x$, 且 $b_n \overset{\text{def.}}{=\!=} \left(\sum\limits_{k=1}^{n} a_k\right) \uparrow \infty$, 则取

$$a_{n,k} = \frac{a_k}{b_n} \quad (k = 1, 2, \cdots, k_n, k_n = n),$$

必有

$$\sum_{k=1}^{k_n} a_{n,k} \to 1, \quad x_n \to x \quad (\text{当 } n \to \infty),$$

所以

$$\frac{1}{b_n}\sum_{k=1}^{n}a_kx_k=\sum_{k=1}^{k_n}a_{n,k}x_k=x'_k\to x.$$

引理 5.3 证毕.

**推论 5.1** 若 $\sum\limits_{k=1}^{\infty}a_k$ 收敛, 且 $b_n\uparrow\infty$, 则

$$\frac{b_1a_1+\cdots+b_na_n}{b_n}\to 0. \tag{5.15}$$

**证** 令 $S_0=0, S_n=a_1+\cdots+a_n$ $(n\geqslant 1)$, 则

$$\begin{aligned}\frac{1}{b_n}\sum_{k=1}^{n}a_kb_k&=\frac{1}{b_n}\sum_{k=1}^{n}(S_k-S_{k-1})b_k\\&=\left(S_n-\frac{1}{b_n}\sum_{k=0}^{n-1}(b_{k+1}-b_k)S_k\right)\to 0.\end{aligned}$$

**定理 5.4 (强大数定律)** 设 $\{X_n,n\geqslant 1\}$ 是相互独立随机变量序列, $\{b_n, n\geqslant 0\}$ 是单调上升趋于 $\infty$ 的正实数列. 若 $\boldsymbol{E}(X_n)=p_n$ 和 $\mathrm{var}(X_n)=\sigma_n^2$ 存在 $(n\geqslant 1)$, 且 $\sum\limits_{n=1}^{\infty}\dfrac{\sigma_n^2}{b_n^2}<\infty$, 则

$$\lim_{n\to\infty}\frac{1}{b_n}\sum_{k=1}^{n}(X_k-p_k)=0,\ \text{a.s.}, \tag{5.16}$$

且

$$\frac{1}{b_n}\sum_{k=1}^{n}(X_k-p_k)\xrightarrow{L^2}0. \tag{5.17}$$

**证** 应用定理 5.2 于 $\left\{\dfrac{X_n-p_n}{b_n},\ n\geqslant 1\right\}$ 有

$$\sum_{n=1}^{\infty}\frac{X_n-p_n}{b_n}=X,\ \text{a.s.}.$$

取 $a_n=\dfrac{1}{b_n}(X_n-p_n)$, 再应用推论 5.1 可得 (5.16).

又因为

$$\boldsymbol{E}\left(\left(\frac{1}{b_n}\sum_{k=1}^{n}(X_k-p_k)\right)^2\right)=\frac{1}{b_n^2}\sum_{k=1}^{n}\sigma_k^2, \tag{5.18}$$

而

$$\sum_{n=1}^{\infty}\frac{\sigma_n^2}{b_n^2}<\infty,$$

所以, 再一次应用推论 5.1 可得

$$\frac{1}{b_n^2}\sum_{k=1}^{n}\sigma_k^2 \to 0. \tag{5.19}$$

由 (5.18)、(5.19) 可得 (5.17).

**引理 5.4** 对任何分布函数 $F(x)$, 恒有

$$\int_{\mathbf{R}}|x|F(dx)<\infty \Leftrightarrow \sum_{n=1}^{\infty}\int_{|x|\geqslant n}F(dx)<\infty. \tag{5.20}$$

**证** 令

$$\varphi(x)=\int_{|y|\geqslant x}F(dy),$$

则

$$\varphi(n)\leqslant\int_{n-1}^{n}\varphi(x)dx\leqslant\varphi(n-1)\quad(\forall n\geqslant 1).$$

所以

$$\begin{aligned}\sum_{n=1}^{\infty}\varphi(n)\leqslant\int_{0}^{\infty}\varphi(x)dx&\leqslant\varphi(0)+\sum_{n=1}^{\infty}\varphi(n)\\&=1+\sum_{n=1}^{\infty}\varphi(n),\end{aligned}$$

因此, $\sum\limits_{n=1}^{\infty}\varphi(n)<\infty$ 的充分必要条件是 $\int_0^\infty\varphi(x)dx<\infty$. 但是

$$\begin{aligned}\int_0^\infty\varphi(x)dx&=\int_0^\infty\left(\int_{|y|\geqslant x}F(dy)\right)dx=\int_{\mathbf{R}}\left(\int_0^{|y|}dx\right)F(dy)\\&=\int_{\mathbf{R}}|y|F(dy).\end{aligned}$$

引理 5.4 证毕.

**定理 5.5 (Kolmogorov 强大数定律)** 设 $\{X_n,n\geqslant 1\}$ 是一列相互独立的具有公共分布函数 $F(x)$ 的随机变量. 若存在随机变量 $X$, 使

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k=X,\quad \text{a.s.}, \tag{5.21}$$

则 $X_k$ 的数学期望存在; 反之若 $X_k$ 的数学期望存在, 则

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}X_k=\boldsymbol{E}(X_1),\quad \text{a.s.}.$$

**证** 令

$$A_n = \left\{ \left| \frac{X_n}{n} \right| \geqslant 1 \right\}.$$

若 (5.21) 成立, 则

$$\frac{X_n}{n} = \left[ \frac{1}{n} \sum_{k=1}^{n} X_k - \frac{n-1}{n} \left( \frac{1}{n-1} \sum_{k=1}^{n-1} X_k \right) \right] \to 0, \quad \text{a.s..}$$

所以

$$P\left( \lim_{n\to\infty} \sup A_n \right) = 0.$$

由 Borel – Contelli 引理可知

$$\sum_{n=1}^{\infty} P(A_n) < \infty. \tag{5.22}$$

故由引理 5.4 及 (5.22) 式知 $X_k$ 的数学期望存在.

反之, 若 $X_k$ 的数学期望存在, 即

$$\int_{\mathbf{R}} |x| F(dx) < \infty, \tag{5.23}$$

作两列随机变量如下:

$$U_n = \begin{cases} X_n, & \text{当 } |X_n| < n, \\ 0, & \text{当 } |X_n| \geqslant n; \end{cases} \quad V_n = \begin{cases} 0, & \text{当 } |X_n| < n, \\ X_n, & \text{当 } |X_n| \geqslant n, \end{cases}$$

则 $\{U_n, n \geqslant 1\}$ 和 $\{V_n, n \geqslant 1\}$ 是两列各自内部相互独立的两列随机变量而且 $X_n = U_n + V_n$ $(\forall n \geqslant 1)$. 令 $\sigma_n^2 = \operatorname{var}(U_n)$, 则

$$\begin{aligned} \sigma_k^2 \leqslant E(U_k^2) &= \int_{|x|<k} x^2 F(dx) \\ &= \sum_{m=1}^{k} \int_{m-1 \leqslant |x| < m} x^2 F(dx) \\ &\leqslant \sum_{m=1}^{k} m \int_{m-1 \leqslant |x| < m} |x| F(dx). \end{aligned} \tag{5.24}$$

令

$$a_m = \int_{m-1 \leqslant |x| < m} |x| F(dx),$$

则

$$\begin{aligned}
\sum_{k=1}^{\infty}\frac{\sigma_k^2}{k^2} &\leqslant \sum_{k=1}^{\infty}\frac{1}{k^2}\sum_{m=1}^{k} m\, a_m \\
&= \sum_{k=1}^{\infty} m\, a_m \sum_{k=m}^{\infty}\frac{1}{k^2} \leqslant \sum_{m=1}^{\infty} m\, a_m\left(\frac{1}{m^2}+\int_m^{\infty}\frac{dx}{x^2}\right) \\
&\leqslant \sum_{m=1}^{\infty} m\, a_m\left(\frac{1}{m^2}+\frac{1}{m}\right) \leqslant 2\sum_{m=1}^{\infty} a_m \\
&= 2\boldsymbol{E}(|X_k|) < \infty.
\end{aligned}$$

所以, 由定理 5.4 得

$$\frac{1}{n}\sum_{k=1}^{n}(U_k-\boldsymbol{E}(U_k)) \to 0 \quad (\text{当 } n\to\infty), \quad \text{a.s.}, \tag{5.25}$$

但是

$$\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{E}(U_k) \to \boldsymbol{E}(X_1). \tag{5.26}$$

由 (5.25) 和 (5.26) 得:

$$\frac{1}{n}\sum_{k=1}^{n}(U_k-\boldsymbol{E}(X_1)) \to 0, \quad \text{a.s..} \tag{5.27}$$

若还能证 $V_n \to 0$, a.s., 则由 (5.27) 和 $U_n$ 及 $V_n$ 之定义可推知

$$\frac{1}{n}\sum_{k=1}^{n}X_k = \frac{1}{n}\left(\sum_{k=1}^{n}U_k+\sum_{k=1}^{n}V_k\right) \to \boldsymbol{E}(X_1), \quad \text{a.s.},$$

即是定理获证. 事实上, 若令 $E_n = \{V_n \neq 0\}$, 则

$$\begin{aligned}
P(E_n) &= \int_{|x|\geqslant n} F(dx) = \sum_{m=1}^{\infty}\int_{n+m-1\leqslant|x|<n+m} F(dx) \\
&\leqslant \sum_{m=1}^{\infty}\frac{1}{n+m-1}\bullet a_{n+m} = \sum_{m=n}^{\infty}\frac{a_{m+1}}{m}.
\end{aligned}$$

所以

$$\sum_{n=1}^{\infty}P(E_n) \leqslant \sum_{n=1}^{\infty}\sum_{m=n}^{\infty}\frac{a_{m+1}}{m} = \sum_{m=1}^{\infty}a_{m+1} < \infty.$$

因此, 由 Borel – Contelli 引理得

$$P\left(\limsup_{n\to\infty} E_n\right) = 0.$$

所以, $V_n \to 0$, a.s.. 定理证毕.

**定理 5.6** 设 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, $F_n(x)$ 是 $X_n$ 的分布函数, 则下列陈述等价:

(1) 存在随机变量 $X$ 使

$$\sum_{n=1}^{\infty} X_n = X, \quad \text{a.s.};$$

(2) 对一切正数序列 $\{a_n, n \geqslant 1\}$ 和 $\{b_n, n \geqslant 1\}$, 只要

$$0 < \liminf_{n\to\infty} a_n, \quad \limsup_{n\to\infty} a_n < \infty,$$
$$0 < \liminf_{n\to\infty} b_n, \quad \limsup_{n\to\infty} b_n < \infty,$$

则有

$$\sum_{n=1}^{\infty} \int_{\mathbf{R}-(-a_n,b_n)} F_n(dx) < \infty, \tag{5.28}$$

$$\sum_{n=1}^{\infty} \int_{(-a_n,b_n)} xF_n(dx)\text{收敛}, \tag{5.29}$$

$$\sum_{n=1}^{\infty} \left( \int_{(-a_n,b_n)} x^2 F_n(dx) - \left[ \int_{(-a_n,b_n)} xF_n(dx) \right]^2 \right) < \infty; \tag{5.30}$$

(3) 对一切正数 $c$, 都有

$$\sum_{n=1}^{\infty} \int_{|x|\geqslant c} F_n(dx) < \infty, \tag{5.31}$$

$$\sum_{n=1}^{\infty} \int_{|x|<c} xF_n(dx) \quad \text{收敛}, \tag{5.32}$$

$$\sum_{n=1}^{\infty} \left( \int_{|x|<c} x^2 F_n(dx) - \left[ \int_{|x|<c} xF_n(dx) \right]^2 \right) < \infty; \tag{5.33}$$

(4) 对某一个正数 $c$, (5.31)、(5.32) 和 (5.33) 成立;

(5) 存在相互独立的随机变量序列 $\{X'_n, n \geqslant 1\}$ 使

$$\boldsymbol{E}((X'_n)^2) < \infty, \quad \boldsymbol{E}(X'_n) = p_n, \quad \operatorname{var}(X'_n) = \sigma_n^2,$$
$$\sum_{n=1}^{\infty} p_n \text{ 收敛}, \quad \sum_{n=1}^{\infty} \sigma_n^2 < \infty, \quad \sum_{n=1}^{\infty} P(X_n \neq X'_n) < \infty.$$

**证** (1) ⇒ (2). 设 (1) 成立且 $\{a_n\}$ 和 $\{b_n\}$ 满足 (2) 中的条件. 令

$$Z_n = \begin{cases} X_n, & \text{当 } -a_n < X_n < b_n, \\ 0, & \text{反之}, \end{cases}$$

再令 $E_n = \{X_n \neq Z_n\}$. 由于 (1) 成立, 所以

$$X_n \to 0, \quad \text{a.s.}.$$

因此

$$P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{X_k = Z_k\}\right) = 1, \tag{5.34}$$

从而存在随机变量 $Z$ 使

$$\sum_{n=1}^{\infty} Z_n = Z, \quad \text{a.s.}.$$

又因为 $Z_n$ 一致 (对 $n$ 来说) 有界, 所以应用定理 5.3 可知

$$\sum_{n=1}^{\infty} \boldsymbol{E}(Z_n) \text{ 和 } \sum_{n=1}^{\infty} \operatorname{var}(Z_n) \text{ 都收敛.}$$

此即 (2) 中的 (5.29) 和 (5.30) 都成立. 而

$$\begin{aligned} P\left(\limsup_{n\to\infty} E_n\right) &= P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} E_k\right) \\ &= 1 - P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} E_k^{\mathrm{c}}\right) \\ &= 1 - P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{X_k = Z_k\}\right) \overset{(5.34)}{=\!=\!=} 0, \end{aligned}$$

所以由 Borel – Contelli 引理知

$$\sum_{n=1}^{\infty} P(E_n) < \infty,$$

此即 (5.28) 成立.

(2) ⇒ (3) ⇒ (4). 这是显然的.

(4) ⇒ (5). 设 (4) 成立. 取

$$X_n' = \begin{cases} X_n, & \text{当 } |X_n| < C, \\ 0, & \text{反之,} \end{cases}$$

则 $\{X_n'\}$ 即为 (5) 中所求.

(5) ⇒ (1). 设 (5) 成立. 则由定理 5.2 知存在 $X'$ 使

$$\sum_{n=1}^{\infty} X_n' = X', \quad \text{a.s.}. \tag{5.35}$$

而

$$\sum_{n=1}^{\infty} P(X_n \neq X_n') < \infty,$$

所以由 Borel - Contelli 引理得

$$P\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{X_k \neq X_k'\}\right) = 0,$$

故

$$P\left(\bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \{X_k = X_k'\}\right) = 1. \tag{5.36}$$

由 (5.35) 和 (5.36) 知: 存在随机变量 $X$ 使

$$\sum_{n=1}^{\infty} X_n = X, \quad \text{a.s..}$$

定理证毕.

在第三章中, 我们曾经证明过随机变量序列 $\{Y_n, n \geqslant 1\}$ 的 a.s. 收敛蕴涵了 $P$ 收敛, 而 $P$ 收敛又蕴涵了 $Y_n$ 的分布的弱收敛. 但这些蕴涵关系的逆命题一般来说是不成立的. 下面我们要证明: 如果 $Y_n = \sum\limits_{k=1}^{n} X_k, n \geqslant 1$, $\{X_k, k \geqslant 1\}$ 是相互独立的随机变量序列, 则前述蕴涵关系的逆命题也成立.

**定义 5.1** 称无穷乘积 $\prod\limits_{n=1}^{\infty} a_n$ 收敛, 如果存在一个 $n_0$, 使

$$\lim_{n\to\infty} a_{n_0+1}a_{n_0+2}\cdots a_n$$

存在而且不等于 0.

称函数项的无穷乘积 $\prod\limits_{n=1}^{\infty} Z(\lambda)(\lambda \in A)$ 依强控意义下收敛, 如果存在一个正项的收敛的数值级数 $\sum\limits_{n=1}^{\infty} C_n$, 使

$$\sum_{n=1}^{\infty} |Z_n(\lambda) - 1| \ll \sum_{n=1}^{\infty} C_n \quad (\forall \lambda \in A).$$

(所谓 $\sum\limits_{n=1}^{\infty} a_n \ll \sum\limits_{n=1}^{\infty} b_n$, 意即对一切 $n \geqslant 1$, 都有 $a_n \leqslant b_n$.)

**定理 5.7** 设 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, $F_n(x)$ 和 $f_n(t)$ 分别为 $X_n$ 的分布函数和特征函数, $\mu_n$ 是 $X_n$ 的中位数, $\tau$ 是任意给定的一个正数. 令

$$\begin{aligned} b_n &= \mu_n + \int_{|x-\mu_n|<\tau} (x - \mu_n) F_n(dx) \\ &= \mu_n + \int_{|x|<\tau} x F_n^{\mu}(dx), \end{aligned}$$

其中 $F_n^{\mu}(x)=F_n(x+\mu_n)$ 是 $X_n-\mu_n$ 的分布函数, 则下列陈述等价:

(1) $\lim\limits_{n\to\infty}\prod\limits_{k=1}^{n}|f_k(t)|^2=f(t)\quad(t\in\mathbf{R})$, $f(t)$ 是特征函数;

(2) $\sum\limits_{n=1}^{\infty}(X_n-b_n)$ a.s. 收敛;

(3) 存在实数列 $\{C_n,n\geqslant 1\}$ 使

$$\sum_{n=1}^{\infty}(X_n-C_n)\quad \text{a.s. 收敛;}$$

(4) $\prod\limits_{n=1}^{\infty}|f_n(t)|^2$ 在 $t\in\mathbf{R}$ 收敛;

(5) $\prod\limits_{n=1}^{\infty}|f_n(t)|^2$ 在 $t\in A$ 收敛,

其中 $A$ 具有 Lebesgue 正测度 $L(A)>0$;

(6) 在每一个有限 $t$ 区间 $[-T,T]$ 上,

$$\prod_{n=1}^{\infty}e^{-itb_n}f_n(t)$$

依强控意义收敛.

**证** (1) $\Rightarrow$ (2). 设

$$\lim_{n\to\infty}\prod_{k=1}^{n}|f_k(t)|^2=f(t)$$

是特征函数, 则可取 $\delta>0$, 使得当 $|t|\leqslant\delta$ 时有 $f(t)\geqslant\dfrac{1}{2}$. (注意 $f(t)$ 是实值特征函数, 当然在 0 点连续, 又有 $f(0)=1$. 所以上述 $\delta$ 存在.) 所以

$$-\sum_{n=1}^{\infty}\log|f_n(t)|^2=-\log f(t)\leqslant\log 2\quad(\text{当 }|t|\leqslant\delta).$$

但是

$$x\leqslant-\log(1-x)\quad(\text{当 }0\leqslant x\leqslant 1\text{ 时}),$$

所以

$$\begin{aligned}\sum_{n=1}^{\infty}(1-|f_n(t)|^2)&\leqslant-\sum_{n=1}^{\infty}\log|f_n(t)|^2\\&\leqslant\log 2\quad(\text{当 }|t|\leqslant\delta\text{ 时}).\end{aligned}\tag{5.37}$$

令

$$X_n'=\begin{cases}X_n, & \text{当 }|X_n-\mu_n|<\tau,\\ \mu_n, & \text{反之,}\end{cases}$$

则可算出 $\boldsymbol{E}(X_n') = b_n$, 且

$$\operatorname{var}(X_n') = \int_{|x-\mu_n|<\tau} (x-\mu_n)^2 F_n(dx) - \left( \int_{|x-\mu_n|<\tau} (x-\mu_n) F_n(dx) \right)^2$$

利用第四章截尾不等式 (2.30), 可得

$$\operatorname{var}(X_n') \leqslant C(\tau, \delta) \int_0^\delta (1 - |f_n(t)|^2) dt. \tag{5.38}$$

由 (5.37)、(5.38) 可得

$$\sum_{n=1}^\infty \operatorname{var}(X_n') \leqslant C(\tau, \delta) \delta \log 2 < \infty,$$

所以, 用定理 5.2 于 $\{X_n' - b_n, n \geqslant 1\}$ 则得

$$\sum_{n=1}^\infty (X_n' - b_n) \quad \text{a.s. 收敛}. \tag{5.39}$$

但是

$$\sum_{n=1}^\infty P(X_n \neq X_n') \leqslant \sum_{n=1}^\infty P(|X_n - \mu_n| \geqslant \tau),$$

利用第四章 (2.27) 可得

$$P(|X_n - \mu_n| \geqslant \tau) = \int_{|x-\mu_n|\geqslant\tau} F_n(dx) \leqslant C_1(\tau, \delta) \int_0^\delta (1 - |f_n(t)|^2) dt,$$

所以

$$\sum_{n=1}^\infty P(X_n \neq X_n') \leqslant C_1(\tau, \delta) \sum_{n=1}^\infty \int_0^\delta (1 - |f_n(t)|^2) dt \leqslant C_1(\tau, \delta) \delta \log 2 < \infty, \tag{5.40}$$

由 (5.40)、(5.39) 和 Borel – Contelli 引理得知:

$$\sum_{n=1}^\infty (X_n - b_n) \quad \text{a.s. 收敛}.$$

(2) ⇒ (3). 显然成立.

(3) ⇒ (4). 设

$$\sum_{n=1}^{\infty}(X_n - C_n) = S, \quad \text{a.s.}, \quad \sum_{n=m}^{\infty}(X_n - C_n) = S_m, \quad \text{a.s.},$$

则 $S_m$ 的特征函数为

$$\lim_{n\to\infty}\prod_{k=m}^{n} e^{-iC_k t} f_k(t) = g_m(t).$$

又因为 $\lim\limits_{m\to\infty} S_m = 0$, a.s., 所以 $\lim\limits_{m\to\infty} g_m(t) = 1$. 因此

$$\lim_{m\to\infty}\lim_{n\to\infty}|f_m(t)|^2\cdots|f_n(t)|^2 = \lim_{m\to\infty}|g_m(t)|^2 = 1.$$

所以,

$$\prod_{n=1}^{\infty}|f_n(t)|^2 \ \text{收敛} \quad (t \in \mathbf{R}).$$

(4) ⇒ (5). 显然成立.

(5) ⇒ (6). 设 (5) 成立, 即存在一个 Lebesgue 正测集 $A$, $L(A) > 0$, 使

$$\lim_{m\to\infty}\lim_{n\to\infty}|f_{m+1}(t)|^2\cdots|f_n(t)|^2 = 1 \ (\forall t \in A).$$

所以存在 $B \subset A, L(B) > 0$, 使

$$\lim_{m\to\infty}\lim_{n\to\infty}|f_{m+1}(t)|^2\cdots|f_n(t)|^2 = 1 \ \text{对}\ t \in B \ \text{一致成立}.$$

由于 $L(B) > 0$, 所以 $B$ 在 $[0,\infty)$ 内或者在 $(-\infty, 0)$ 内有正测度子集, 而 $|f_k(t)|^2$ 是偶函数, 所以不失普遍性可设 $B\bigcap[0,\infty)$ 有正测度. 因此, 存在 $\delta > 0$, 使 $B\bigcap(0,\delta) \stackrel{\text{def.}}{=\!=} D$ 有正测度 $\rho$. 取 $m_0$ 使

$$\lim_{n\to\infty}|f_{m_0+1}(t)|^2\cdots|f_n(t)|^2 \geqslant \frac{1}{2} \quad (\forall t \in D).$$

仿 (5.37) 可证

$$\sum_{n=m_0+1}^{\infty}|f_n(t)|^2 \leqslant \log 2 \quad (\forall t \in D).$$

利用第四章不等式 (2.24) 得:

$$|e^{-ib_n t} f_n(t) - 1| \leqslant L_0(T, \tau, \rho, \delta)\int_D (1 - |f_n(t)|^2)dt, \quad (|t| \leqslant T).$$

取 $C_n = L_0(T,\tau,\rho,\delta)\int_D(1 - |f_n(t)|^2)dt$, 则 $\sum\limits_{n=1}^{\infty} C_n$ 收敛且

$$\sum_{n=1}^{\infty}|e^{-ib_n t} f_n(t) - 1| \ll \sum_{n=1}^{\infty} C_n \quad (\text{当}\ |t| \leqslant T),$$

此即 (6) 成立.

(6)⇒ (1). 设 (6) 成立, 亦即 $\prod\limits_{n=1}^{\infty} e^{-ib_nt}f_n(t)$ 在任意有限 $t$ 区间 $[-T,T]$ 上依强控意义下收敛, 则

$$\prod_{n=1}^{\infty} e^{-ib_nt}f_n(t)$$

在 $|t| \leqslant T$ 上一致收敛, 因此 $\prod\limits_{n=1}^{\infty} |f_n(t)|$ 在 $|t| \leqslant T$ 上一致收敛. 由于 $T$ 可为任意正数, 所以 $\prod\limits_{n=1}^{\infty} |f_n(t)|^2$ 的极限函数是特征函数, 此即 (1) 成立. 定理证毕.

**定理 5.8** 设 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, $f_n(t)$ 是 $X_n$ 的特征函数, 则下列陈述等价:

(1) $\lim\limits_{n\to\infty} f_1(t)\cdots f_n(t) = f(t)$ 是特征函数;

(2) $\sum\limits_{n=1}^{\infty} X_n$ a.s. 收敛;

(3) $\prod\limits_{n=1}^{\infty} f_n(t)$ 收敛 $(\forall t \in \mathbf{R})$;

(4) $\prod\limits_{n=1}^{\infty} f_n(t)$ 收敛 $(\forall t \in E, L(E) > 0)$.

**证** (1)⇒(2). 设 (1) 成立, 则

$$\lim_{n\to\infty} |f_1(t)|^2 \cdots |f_n(t)|^2 = |f(t)|^2,$$

所以由定理 5.7 知

$$\sum_{n=1}^{\infty}(X_n - b_n) \quad \text{a.s. 收敛}.$$

故

$$\lim_{n\to\infty} e^{-ib_1t}f_1(t)\cdots e^{-ib_nt}f_n(t) = g(t), \tag{5.41}$$

是特征函数. 所以存在一个 $\delta > 0$, 使 $|t| < \delta$ 时 $|g(t)| > 0$. 因此, 由 (5.41) 知

$$\lim_{n\to\infty} e^{-i(b_1+\cdots+b_n)t} = h(t) \quad (\text{当 } |t| < \delta).$$

由上式易证 (参见 [36] 第五章引理 1.5):

$$\sum_{n=1}^{\infty} b_n \quad \text{收敛到有限数}.$$

但是 $\lim\limits_{n\to\infty} e^{-i(b_1+\cdots+b_n)t}f_1(t)\cdots f_n(t) = \varphi(t)$ 是特征函数, (因为 $\prod\limits_{n=1}^{\infty} e^{-ib_nt}f_n(t)$ 依强控意义收敛) 所以

$$\lim_{n\to\infty} f_1(t)\cdots f_n(t) = e^{i\sum\limits_{n=1}^{\infty} b_nt}\varphi(t)$$

是特征函数. 定理证毕.

**推论 5.1** 设 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, $\{b_n, n \geqslant 1\}$ 如定理 5.7 所定义. 如果 $\sum_{n=1}^{\infty} X_n$ a.s. 收敛, 则 $\sum_{n=1}^{\infty} b_n$ 收敛到有限数.

**推论 5.2** 设 $\{X_n, n \geqslant 1\}$ 是相互独立的随机变量序列, $f_n(t)$ 是 $X_n$ 的特征函数, 则下列陈述等价:

(1) $\sum_{k=1}^{n} X_k \xrightarrow{\text{a.s.}} X$;

(2) $\sum_{k=1}^{n} X_k \xrightarrow{P} X$;

(3) $\prod_{k=1}^{n} f_k(t) \to f(t)$ 是特征函数.

其中 $f(t)$ 是 $X$ 的特征函数.

## *§6 重对数律

在第 5 节中, Kolmogorov 强大数定律说, 对于相互独立相同分布的随机变量序列 $\{X_n, n \geqslant 1\}$ 来说, 它服从强大数定律的充分必要条件是 $X_n$ 的数学期望存在. 特别地, 当 $\boldsymbol{E}(X_n) = 0$ 时, 若令 $S_n = \sum_{k=1}^{n} X_k$, 则

$$\lim_{n\to\infty} \frac{S_n}{n} = 0, \quad \text{a.s.}. \tag{6.1}$$

现在, 我们要讨论较强大数定律更细致的形式: 对于什么样的正实数列 $\{c_n, n \geqslant 1\}$, 可保证

$$\limsup_{n\to\infty} \frac{S_n}{c_n} = 1, \quad \text{a.s.}.$$

这就是重对数律问题. 第一个较一般的重对数律是 Kolmogorov 1929 年给出的.

**定义 6.1** 对于集合列 $\{A_n, n \geqslant 1\}$, 记

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_n = \{A_n, \text{i.o.}\}.$$

设 $\{S_n, n \geqslant 1\}$ 是一列随机变量, $\{b_n, n \geqslant 1\}$ 是一列实数, 称 $\{b_n\}$ 属于 $\{S_n\}$ 的上 (相应地, 下) 类, 如果

$$P(S_n > b_n, \text{i.o.}) = 0 \quad (\text{相应地}, 1).$$

在证明重对数律以前, 我们研究尾概率的指数不等式.

**引理 6.1**　设 $X$ 是有界随机变量: $|X| \leqslant c$, 且 $\boldsymbol{E}(X)=0, \sigma^2=\operatorname{var}(X), t>0$. 若 $tc \leqslant 1$, 则

$$\exp\left[\frac{t^2\sigma^2}{2}(1-tc)\right] < \boldsymbol{E}(e^{tX}) < \exp\left[\frac{t^2\sigma^2}{2}\left(1+\frac{tc}{2}\right)\right]. \tag{6.2}$$

**证**　因为 $|X| \leqslant c, \boldsymbol{E}(X)=0$, 所以

$$\begin{aligned} &|\boldsymbol{E}(X^n)| \leqslant c^n, \\ &\boldsymbol{E}(e^{tX}) = 1+\frac{t^2}{2!}\boldsymbol{E}(X^2)+\frac{t^3}{3!}\boldsymbol{E}(X^3)+\cdots. \end{aligned} \tag{6.3}$$

又因为

$$e^{t(1-t)} < 1+t < e^t, \tag{6.4}$$

所以由 (6.3) 和 (6.4) 及 $tc \leqslant 1$ 得:

$$\begin{aligned} \boldsymbol{E}(e^{tX}) &< 1+\frac{t^2\sigma^2}{2}\left(1+\frac{tc}{3}+\frac{t^2c^2}{3\cdot 4}+\cdots\right) \\ &< 1+\frac{t^2\sigma^2}{2}\left(1+\frac{tc}{2}\right) < \exp\left[\frac{t^2\sigma^2}{2}\left(1+\frac{tc}{2}\right)\right], \end{aligned} \tag{6.5}$$

且

$$\begin{aligned} \boldsymbol{E}(e^{tX}) &> 1+\frac{t^2\sigma^2}{2}\left(1-\frac{tc}{3}-\frac{t^2c^2}{3\cdot 4}-\cdots\right) \\ &> 1+\frac{t^2\sigma^2}{2}\left(1-\frac{tc}{2}\right) > \exp\left[\frac{t^2\sigma^2}{2}(1-tc)\right]. \end{aligned} \tag{6.6}$$

由 (6.5)、(6.6) 即得 (6.2). 引理证毕.

**定理 6.1**　设 $\{X_n, n \geqslant 1\}$ 是相互独立的数学期望为 0 且具有有限方差的随机变量序列. 令

$$S_n = \sum_{k=1}^n X_k, \quad s_n^2 = \operatorname{var}(S_n) = \sum_{k=1}^n \operatorname{var}(X_k),$$

$$c = \sup_{\substack{1\leqslant k\leqslant n \\ \omega\in\Omega}} \left|\frac{X_k(\omega)}{s_n}\right|, \quad \varepsilon > 0, \text{则}$$

(1) 当 $\varepsilon c \leqslant 1$ 时有

$$P\left(\frac{S_n}{s_n} > \varepsilon\right) < \exp\left\{-\frac{\varepsilon^2}{2}\left(1-\frac{\varepsilon c}{2}\right)\right\}; \tag{6.7-1}$$

而当 $\varepsilon c \geqslant 1$ 时有

$$P\left(\frac{S_n}{s_n} > \varepsilon\right) < \exp\left\{-\frac{\varepsilon}{4c}\right\}. \tag{6.7-2}$$

(2) 给定 $\gamma>0$, 当 $c=c(\gamma)$ 充分小, $\varepsilon=\varepsilon(\gamma)$ 充分大时有

$$P\left(\frac{S_n}{s_n}>\varepsilon\right)>\exp\left\{-\frac{\varepsilon^2}{2}(1+\gamma)\right\}. \tag{6.8}$$

**证** (1) 令 $S'_n=S_n/s_n$, 并取引理 6.1 中的 $X$ 为 $X_k/s_n$, 再注意

$$\boldsymbol{E}(e^{tS'_n})=\prod_{k=1}^{n}\boldsymbol{E}\left(\exp\left[\frac{tX_k}{s_n}\right]\right), \tag{6.9}$$

则由引理 6.1 知: 当 $tc\leqslant 1$ 时有

$$\exp\left[\frac{t^2}{2}(1-tc)\right]<\boldsymbol{E}(e^{tS'_n})<\exp\left[\frac{t^2}{2}\left(1+\frac{tc}{2}\right)\right]. \tag{6.10}$$

但是

$$\begin{aligned}P(S'_n>\varepsilon)&\leqslant e^{-t\varepsilon}\boldsymbol{E}(e^{tS'_n})\\&\leqslant\exp\left[-t\varepsilon+\frac{t^2}{2}\left(1+\frac{tc}{2}\right)\right].\end{aligned} \tag{6.11}$$

当 $\varepsilon c\leqslant 1$ 时在 (6.11) 中取 $t=\varepsilon$ 即得 (6.7 − 1).

当 $\varepsilon c\geqslant 1$ 时, 取 $t=\dfrac{1}{c}$, 引理 6.1 的条件仍然满足, (6.11) 仍然成立, 而 (6.11) 中的 $t$ 代之以 $\dfrac{1}{c}$ 后即得

$$\begin{aligned}P(S'_n>\varepsilon)&\leqslant\exp\left[-\frac{\varepsilon}{c}+\frac{1}{2c^2}\left(1+\frac{1}{2}\right)\right]\\&\leqslant\exp\left[-\frac{\varepsilon}{c}+\frac{3\varepsilon}{4c}\right]\\&=\exp\left[-\frac{\varepsilon}{4c}\right],\end{aligned}$$

此即 (6.7 − 2) 成立.

(2) 之证明类似于 (1). 详细推导请见 [90] p.267 − 269.

**定理 6.2 (Kolmogorov 重对数律)** 设 $\{X_n,n\geqslant 1\}$ 是一列相互独立的数学期望为 0 方差有限的随机变量. 令 $S_n=\sum\limits_{k=1}^{n}X_k, s_n^2=\operatorname{var}(S_n), t_n=(2\log\log s_n^2)^{\frac{1}{2}}$. 若

$$\lim_{n\to\infty}s_n^2=\infty,\quad |X_n|\leqslant M_n=o\left(\left[\frac{s_n^2}{\log\log s_n^2}\right]^{\frac{1}{2}}\right), \tag{6.12}$$

其中 $\{M_n,n\geqslant 1\}$ 是一列正的常数, 则

$$P\left(\limsup_{n\to\infty}\frac{S_n}{s_nt_n}=1\right)=1. \tag{6.13}$$

**证** 由假设 (6.12) 知

$$s_n^2 \to \infty, \quad \frac{s_{n+1}^2}{s_n^2} \to 1 \quad (\text{当 } n \to \infty \text{ 时}).$$

所以, 任给常数 $c > 1$, 存在单增的趋于 $\infty$ 的正整数列 $\{n_k = n_k(c), k \geqslant 1\}$, 使

$$s_{n_k} \sim c^k, \tag{6.14}$$

此处 $a_k \sim c^k$ 意即 $\lim\limits_{k\to\infty}(a_k/c^k) = 1$. (取 $n_k = \min\{r : s_r \geqslant c^k, r \text{ 为正整数}\}$ 即可.)

我们分几步证明此定理. 下面的 $\delta, \delta'$ 和 $\delta''$ 都是适当取定的正数.

(1) 首先我们证明 $\{(1+\delta)s_n t_n, n \geqslant 1\}$ 属于 $\{S_n, n \geqslant 1\}$ 的上类.

令

$$S_{n_k}^* = \max_{1\leqslant n\leqslant n_k} S_n.$$

由于 $\{s_n, n \geqslant 1\}$ 和 $\{t_n, n \geqslant 1\}$ 单增, 所以

$$\begin{aligned} &P(\{S_n > (1+\delta)s_n t_n\}, \text{i.o.}) \\ &\leqslant P(\{S_{n_k}^* > (1+\delta)s_{n_{k-1}} t_{n_{k-1}}\}, \text{i.o.}). \end{aligned} \tag{6.15}$$

但是由 $\{n_k\}$ 的取法和 $t_{n_k} = (2\log\log s_{n_k}^2)^2$ 可知:

$$(1+\delta)s_{n_{k-1}} t_{n_{k-1}} \sim \frac{1+\delta}{c} s_{n_k} t_{n_k}, \tag{6.16}$$

因此, 取 $\delta' < \delta$ 时, 可以选 $c > 1$ 使

$$\frac{1+\delta}{c} > 1 + \delta', \tag{6.17}$$

且

$$\begin{aligned} &P(\{S_{n_k}^* > (1+\delta)s_{n_{k-1}} t_{n_{k-1}}\}, \text{i.o.}) \\ &\leqslant P(\{S_{n_k}^* > (1+\delta')s_{n_k} t_{n_k}\}, \text{i.o.}). \end{aligned} \tag{6.18}$$

由 (6.15) 和 (6.18) 和 Borel – Contelli 引理, 为证 $\{(1+\delta)s_n, t_n, n \geqslant 1\}$ 属于 $\{S_n, n \geqslant 1\}$ 的上类, 只需证

$$\sum_{k=1}^{\infty} P\left(\{S_{n_k}^* > (1+\delta')s_{n_k} t_{n_k}\}\right) < \infty. \tag{6.19}$$

由第三章推论 2.3 知

$$\begin{aligned} &P\left(\{S_{n_k}^* > (1+\delta')s_{n_k} t_{n_k}\}\right) \\ &\leqslant 2P\left(\left\{S_{n_k} > \left(1 + \delta'' - \frac{\sqrt{2}}{t_{n_k}}\right) s_{n_k} t_{n_k}\right\}\right). \end{aligned} \tag{6.20}$$

此处

$$1+\delta'-\frac{\sqrt{2}}{t_{n_k}} \to 1+\delta' \quad (\text{当 } k\to\infty),$$

所以可取 $\delta''<\delta'$, 且当 $k$ 充分大时有

$$\begin{aligned}
&P\left(S_{n_k} > \left(1+\delta'-\frac{\sqrt{2}}{t_{n_k}}\right) s_{n_k} t_{n_k}\right)\\
&\leqslant P(S_{n_k} > (1+\delta'')s_{n_k}t_{n_k}).
\end{aligned} \tag{6.21}$$

在定理 6.1 的 (6.7 − 1) 式中取 $n=n_k, \varepsilon=\varepsilon_k=(1+\delta'')t_{n_k}, c=c_k=\sup\limits_{\substack{1\leqslant j\leqslant n_k\\ \omega\in\Omega}}\left|\frac{X_j(\omega)}{s_{n_k}}\right|$, 则由条件 (6.12) 知:

$$c_k\varepsilon_k \leqslant (1+\delta'')\max_{1\leqslant j\leqslant n_k}\left[0\left(\left[\frac{(s_j^2/s_{n_k}^2)2\log\log s_{n_k}^2}{\log\log s_j^2}\right]^{\frac{1}{2}}\right)\right] \to 0 \quad (\text{当 } k\to\infty \text{ 时}).$$

所以当 $k$ 充分大时 $c_k\varepsilon_k\leqslant 1$. 故由定理 6.1 的 (6.7 − 1) 有

$$\begin{aligned}
&P(S_{n_k} > (1+\delta'')s_{n_k}t_{n_k})\\
&\leqslant \exp\left\{-\frac{1}{2}(1+\delta'')^2 t_{n_k}^2\left(1-\frac{\varepsilon_k c_k}{2}\right)\right\}\\
&\leqslant \exp\{-(1+\delta'')\log\log s_{n_k}^2\}\\
&\sim \frac{1}{(2k\log c_k)^{1+\delta''}} \quad (\text{当 } k \text{ 充分大}).
\end{aligned} \tag{6.22}$$

由 (6.20)~(6.22) 得 (6.19), 从而 (1) 得证, 即

$$P(\{S_n > (1+\delta)s_n t_n\}, \text{i.o.}) = 0. \tag{6.23}$$

由于以 $-X_n$ 代 $X_n$ 时定理中一切条件仍成立, 故由 (6.23) 得

$$P(\{|S_n| > (1+\delta)s_n t_n\}, \text{i.o.}) = 0. \tag{6.24}$$

(2) 再证 $\{(1-\delta')s_n t_n, n\geqslant 1\}$ 属于 $\{S_n, n\geqslant 1\}$ 的下类, 此处取 $1>\delta'>\delta$. 令

$$u_k^2 = s_{n_k}^2 - s_{n_{k-1}}^2 \sim s_{n_k}^2\left(1-\frac{1}{c^2}\right), \tag{6.25}$$

$$v_k = (2\log\log u_k^2)^{\frac{1}{2}} \sim (2\log\log s_{n_k}^2)^{\frac{1}{2}} = t_{n_k}, \tag{6.26}$$

此处 $c$ 由 (6.14) 式所决定. 再令

$$A_k = \{S_{n_k} - S_{n_{k-1}} > (1-\delta)u_k v_k\}, \quad \varepsilon_k = (1-\delta)v_k. \tag{6.27}$$

首先, 我们要证明

$$P(A_k, \text{i.o.}) = 1. \tag{6.28}$$

显然 $\{(S_{n_k} - S_{n_{k-1}}), k \geqslant 1\}$ 是相互独立的随机变量列. 由 Borel – Contelli 引理, 为证 (6.28) 只需证明

$$\sum_{k=1}^{\infty} P(A_k) = \infty. \tag{6.29}$$

但是当 $k \to \infty$ 时,

$$\varepsilon_k \to \infty, \ c_k \overset{\text{def.}}{=\!=} \sup_{\substack{n_{k-1}<j\leqslant n_k \\ \omega\in\Omega}} \left(\frac{|X_j(\omega)|}{u_k}\right) \to 0. \tag{6.30}$$

取 (6.8) 中的 $\gamma+1 = \dfrac{1}{1-\delta}$, 若注意 $u_k^2 = \text{var}(S_{n_k} - S_{n_{k-1}})$, 则由 (6.8) 得

$$\begin{aligned} P(A_k) &= P\left(\frac{S_{n_k} - S_{n_{k-1}}}{\sqrt{\text{var}(S_{n_k} - S_{n_{k-1}})}} > (1-\delta)v_k\right) \\ &> \exp\left\{-\frac{1}{2}(1+\gamma)(1-\delta)^2 v_k^2\right\} \\ &= \exp\{-(1-\delta)\log\log u_k^2\} \\ &\sim \frac{1}{(2k\log c)^{1-\delta}} \end{aligned} \tag{6.31}$$

$\left(\text{因为 } u_k^2 \sim s_{n_k}^2\left(1-\dfrac{1}{c^2}\right) \sim c^{2k}\left(1-\dfrac{1}{c^2}\right).\right)$ 所以

$$\sum_{k=1}^{\infty} P(A_k) = \infty,$$

故 (6.28) 成立.

另一方面, 若令 $B_k = \{|S_{n_{k-1}}| \leqslant 2s_{n_{k-1}}t_{n_{k-1}}\}$, 则由 (6.24) 得

$$P(B_k^c, \text{i.o.}) = 0, \quad (B_k^c \text{ 为 } B_k \text{ 之补集}), \tag{6.32}$$

即

$$P\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} B_k\right) = 1. \tag{6.33}$$

但

$$\begin{aligned} P(A_kB_k, \text{i.o.}) &= P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n} A_kB_k\right) \\ &\geqslant P\left(\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right)\bigcap\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} B_k\right)\right). \end{aligned} \tag{6.34}$$

由 (6.28)、(6.33) 和 (6.34) 知

$$P(A_kB_k, \text{i.o}) = 1. \tag{6.35}$$

但是

$$A_kB_k \subset \left\{S_{n_k} > (1-\delta)u_kv_k - 2s_{n_{k-1}}t_{n_{k-1}}\right\}, \tag{6.36}$$

$$(1-\delta)u_kv_k - 2s_{n_{k-1}}t_{n_{k-1}} \sim \left[(1-\delta)\left(1-\frac{1}{c^2}\right)^{\frac{1}{2}} - \frac{2}{c}\right], \tag{6.37}$$

若取 $c$ 充分大, 使 $\delta' > \delta$ 且

$$(1-\delta)\left(1-\frac{1}{c^2}\right)^{\frac{1}{2}} - \frac{2}{c} > 1-\delta',$$

则

$$\begin{aligned} 1 &= P(A_kB_k, \text{i.o.}) \\ &\leqslant P\left(\left\{S_{n_k} > (1-\delta')s_{n_k}t_{n_k}\right\}, \text{i.o.}\right), \end{aligned}$$

更有

$$\begin{aligned} &P\left(\left\{S_n > (1-\delta')s_nt_n\right\}, \text{i.o.}\right) \\ \geqslant &P\left(\left\{S_{n_k} > (1-\delta')s_{n_k}t_{n_k}\right\}, \text{i.o.}\right) = 1. \end{aligned}$$

此即

$\{(1-\delta')s_n, t_n, n \geqslant 1\}$ 属于 $\{S_n, n \geqslant 1\}$ 的下类. 定理证毕.

**推论 6.1** *在定理 6.2 的条件下, 恒有*

$$P\left(\limsup_{n\to\infty} \frac{|S_n|}{s_nt_n} = 1\right) = 1. \tag{6.38}$$

**证** 因为以 $-X_n$ 代 $X_n, (n \geqslant 1)$, 定理 6.2 的条件全部成立. 故 (6.38) 成立.

## §7 习题及应用

1. 设 $\{X_n\}$ 是相互独立相同分布的随机变量序列, 若其分布由

(a) $P(X_n = 2^{k-\log k - 2\log\log k}) = \dfrac{1}{2^k}(k = 1, 2, \cdots)$;

(b) $P(X_n = k) = \dfrac{c}{k^2\log^2 k}\left(k \geqslant 2, \dfrac{1}{c} = \sum\limits_{k=2}^{\infty} \dfrac{1}{k^2\log^2 k}\right)$

所确定, 则 $\{X_n\}$ 服从大数定律.

2. 设 $\{X_n\}$ 是相互独立的随机变量序列, 若其分布由

$$P(X_n = n^\alpha) = P(X_n = -n^\alpha) = \frac{1}{2} \quad (n = 1, 2, \cdots)$$

所确定, 则 $\{X_n\}$ 服从大数定律的充要条件是 $\alpha < \frac{1}{2}$.

3. 设 $\{X_n\}$ 是相互独立的随机变量序列, 若

$$\lim_{A\to\infty}\left(\sup_{n\geqslant 1}\int_{|x|\geqslant A}|x|dF_n(x)\right) = 0,$$

其中 $F_n(x)$ 是 $X_n$ 的 d.f., 则 $\{X_n\}$ 服从大数定律.

4. 设 $\{X_n\}$ 是随机变量序列. 若 $\mathrm{var}(X_n) \leqslant c(n = 1, 2, \cdots)$, 且 $\lim\limits_{|i-j|\to\infty}\rho(X_i, X_j) = 0$, 则 $\{X_n\}$ 服从大数定律. ($\rho(X_i, X_j)$ 为 $X_i$ 与 $X_j$ 的相关系数.)

5. 设 $\{X_n\}$ 是随机变量序列. $\{X_n\}$ 服从大数定律的充要条件是:

$$\lim_{n\to\infty}\boldsymbol{E}\left(\frac{\left[\sum\limits_{k=1}^{n}(X_k - \boldsymbol{E}(X_k))\right]^2}{n^2 + \left[\sum\limits_{k=1}^{n}(X_k - \boldsymbol{E}(X_k))\right]^2}\right) = 0.$$

6. 设 $\{X_n\}$ 为相互独立相同分布的随机变量序列, 而且有有限的数学期望与方差. 令 $S_n = \sum\limits_{k=1}^{n} X_k$, 则 $\{a_n S_n\}$ 服从大数定律 (其中 $a_n$ 是常数, $a_n \to 0$).

7. 设 $\{X_k\}$ 为随机变量序列. 若 $X_k$ 不与 $X_{k-1}$ 和 $X_{k+1}$ 相互独立, 但与其他的 $X_i$ 相互独立, 且 $X_k$ 具有有限方差 $(k = 1, 2, \cdots)$, 则 $\{X_k\}$ 服从大数定律.

8. 试判断下列随机变量序列是否服从中心极限定理:

(a) $P(X_n = -2^n) = P(X_n = 2^n) = \frac{1}{2} \quad (n = 1, 2, \cdots)$;

(b) $P(X_n = -2^n) = P(X_n = 2^n) = 2^{-(2n+1)}$,
$P(X_n = 0) = 1 - 2^{-2n} \quad (n = 1, 2, \cdots)$;

(c) $P(X_n = n) = P(X_n = -n) = \frac{1}{2}n^{-\frac{1}{2}}$,
$P(X_n = 0) = 1 - n^{-\frac{1}{2}} \quad (n = 1, 2, \cdots)$.

9. 设 $\{X_n\}$ 为相互独立的随机变量序列, 而且 $P(X_n = n^\alpha) = P(X_n = -n^\alpha) = \frac{1}{2}\left(\alpha > \frac{1}{2}\right)$, 则 $\{X_n\}$ 服从中心极限定理.

10. 求证:

$$\lim_{n\to\infty} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!} = \frac{1}{2}.$$

11. 设 $S_n$ 是 $n$ 次独立试验中事件 $A$ 出现的次数, $p_k$ 是第 $k$ 次试验中 $A$ 出现的概率, 试证

$$\lim_{n\to\infty} P\left(\frac{S_n-\sum_{k=1}^{n}p_k}{\sqrt{\sum_{k=1}^{n}p_k(1-p_k)}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{t^2}{2}}dt$$

的充要条件是

$$\sum_{k=1}^{\infty}p_k(1-p_k)=\infty.$$

12. 若 $\sum_{k=1}^{\infty}\sigma_k^2/k^2$ 发散, 则存在一串相互独立相同分布的随机变量 $\{X_k\}$, 使 $\mathrm{var}(X_k)=\sigma_k^2$, 而 $\{X_k\}$ 不服从强大数定律 (提示: 首先证明 $\sum_{n=1}^{\infty}P(|X_n|\geqslant\varepsilon n)$ 收敛是 $\{X_k\}$ 服从强大数定律的必要条件).

13. Kolmogorov 不等式的推广. 设 $\{X_n\}$ 是相互独立的随机变量序列, $\boldsymbol{E}(X_n)=0,(n\geqslant1)$. 令 $c_n=\left\{\sup_{k\leqslant n}|S_k|\geqslant c\right\}$, 其中 $S_n=\sum_{k=1}^{n}X_k$, 试证

$$c^rP(C_n)\leqslant\boldsymbol{E}(|S_n|^r\mathbf{1}_{C_n}),$$

其中 $r\geqslant1,\mathbf{1}_{C_n}$ 为 $C_n$ 上的示性函数.

14. 设 $\{X_n\}$ 为相互独立的随机变量序列. 令

$$T_n^r=\sup_{k\leqslant n}|S_k|^r,\quad r\geqslant1,\quad S_n=\sum_{k=1}^{n}X_k.$$

(a) 若 $X_k$ 服从对称分布 $(k\geqslant1)$, 则

$$\boldsymbol{E}(T_n^r)\leqslant2\boldsymbol{E}(|S_n|^r).$$

(b) 若 $\boldsymbol{E}(X_k)=0(k\geqslant1)$, 则

$$\boldsymbol{E}(T_n^r)\leqslant2^{2r+1}\boldsymbol{E}(|S_n|^r).$$

15. 设 $\{X_n\}$ 是相互独立的随机变量序列, $\boldsymbol{E}(X_n)=0(n\geqslant1),r\geqslant1$, 如果

$$\sum_{n=1}^{\infty}\frac{\boldsymbol{E}(|X_n|^{2r})}{n^{r+1}}$$

收敛, 则 $\{X_n\}$ 服从强大数定律.

16. 设 $\{\theta_n\}$ 是相互独立的随机变量序列, 若 $\sum_{n=1}^{\infty}C_n^2<\infty(\{C_n\}$ 是实数列), 则 $\sum_{n=1}^{\infty}C_ne^{i\theta_n}$ 收敛, a.s..

# 第六章　可数状态的 Markov 链

## §1　随机过程的基本概念

17 世纪是概率论的启萌阶段, 概率论并未形成一个有完整理论的数学分支. 人们关注与研究的仅仅是一个 "偶然事件" (或随机事件) $A$ 发生的可能性大小 (概率), 或有限多个随机事件 $A_1, A_2, \cdots, A_n$ 之间的关系. 随着 17、18 世纪古典分析学的发展, 人们才逐渐把随机事件抽象化. 提出随机变量 $X$ 的概念, 而主要关注及研究的是随机变量 $X$ 的概率分布; 有限个随机变量 $X_1, X_2, \cdots, X_n$ 的相互关系和随机变量序列 $X_1, X_2, \cdots$ 的各种极限性质. 由随机变量序列 $\{X_n, n \geqslant 1\}$ 逐渐形成离散时间的随机过程的概念, 进而提出连续时间的随机过程 $\{X_t, t \in [0, \infty)\}$ 的概念. 而真正奠定随机过程的理论基础. 是 Kolmogorov 在 20 世纪 30 年代建立起了 "无穷维乘积空间" 上造测度的理论以后. 这个理论说明了随机过程的存在性, 为随机过程的研究开拓了广阔的领域.

**定义 1.1**　设 $(\Omega, \mathscr{F}, P)$ 是概率空间, $(E, \mathscr{E})$ 为可测空间, $\boldsymbol{T} \subset \mathbf{R}$, 若对任何 $t \in \boldsymbol{T}$,

$$X_t : \Omega \mapsto E$$

且 $X_t \in \mathscr{F}/\mathscr{E}$, 即 $X_t$ 关于 $\mathscr{F}, \mathscr{E}$ 是可测的, 则称 $\{X_t, t \in \boldsymbol{T}\}$ 是 $(\Omega, \mathscr{F}, P)$ 上取值于 $E$ 的**随机过程**, 或 $(E, \mathscr{E})$ 随机过程, 称 $(E, \mathscr{E})$ 为其 "**相空间**" 或 "**状态空间**", 称 $\boldsymbol{T}$ 为其 "**时间域**" 或 "**时间参数集**", 对每个 $\omega \in \Omega, X_\bullet(\omega)$ (从 $\boldsymbol{T}$ 到 $E$ 的映射) 称为相应于 $\omega$ 的 "**轨道**". 实质上, 每个 $X_t$ 都是一个 $E$ 值随机元.

在无混淆的情况下，简称 $\{X_t, t\in \boldsymbol{T}\}$ 为随机过程. 有时记 $X_t = X(t), X_t(\omega) = X(t,\omega), X_\bullet(\omega) = X(\bullet,\omega), X_t(\bullet) = X(t,\bullet)$.

设 $\{\mathscr{F}_t, t\in \boldsymbol{T}\}$ 是 $\mathscr{F}$ 中的一族单增的子 $\sigma$ 代数，即对任何 $t_1, t_2 \in \boldsymbol{T}, t_1 < t_2$，有 $\mathscr{F}_{t_1} \subset \mathscr{F}_{t_2}$. 若对任何 $t\in \boldsymbol{T}, X_t \in \mathscr{F}_t/\mathscr{E}$，则称 $\{X_t, t\in \boldsymbol{T}\}$ 是 $\{\mathscr{F}_t\}$ 适应过程，或称之为适应于 $\{\mathscr{F}_t\}$ 的随机过程.

显然，若 $\mathscr{F}_t = \sigma(X_u, u\leqslant t, u\in \boldsymbol{T})$ 是 $\{X_u, u\leqslant t, u\in \boldsymbol{T}\}$ 所产生的 $\sigma$ 代数，则 $\{X_t, t\in \boldsymbol{T}\}$ 是 $\{\mathscr{F}_t\}$ 适应过程.

本书主要研究的随机过程 $\{X_t, t\in \boldsymbol{T}\}$ 的相空间是 $(\mathbf{R}, \mathscr{B}(\mathbf{R}))$ 或 $(\boldsymbol{B}, \|\bullet\|, \mathscr{B}(\boldsymbol{B}))$，其中 $(\boldsymbol{B}, \|\bullet\|)$ 是 Banach 空间，$\mathscr{B}(\boldsymbol{B})$ 是 $\boldsymbol{B}$ 上的 Borel $\sigma$ 代数，而时间参数集 $\boldsymbol{T} = \{0,1,2,\cdots\}$ 或 $[0,\infty)$ 或 $\{\cdots,-2,-1,0,1,2,\cdots\}$ 或 $(-\infty,\infty)$.

**定义 1.2** 设 $\{X_t, t\in \boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以 $(E,\mathscr{E})$ 为相空间的随机过程，$\boldsymbol{T} = [0,\infty)$，或 $(-\infty,\infty)$，或直线上的任一区间. 称 $\{X_t, t\in \boldsymbol{T}\}$ 是**可测的**，如果对任何 $A\in\mathscr{E}$，都有

$$\{(t,\omega)\in \boldsymbol{T}\times\Omega : X(t,\omega)\in A\} \in \mathscr{B}(\boldsymbol{T})\times\mathscr{F}.$$

若 $\{\mathscr{F}_t, t\in \boldsymbol{T}\}$ 是 $\mathscr{F}$ 中的一族单增的 $\sigma$ 代数，且对任何 $t\in \boldsymbol{T}, A\in\mathscr{E}$ 都有

$$\{(u,\omega)\in [0,t]\times\Omega : X(u,\omega)\in A\} \in \mathscr{B}([0,t])\times\mathscr{F},$$

则称 $\{X_t, t\in \boldsymbol{T}\}$ 关于 $\{\mathscr{F}_t\}$**循序可测**.

**命题 1.1** 设 $\{\mathscr{F}_t, t\in \boldsymbol{T}\}$ 是 $\mathscr{F}$ 中的一族单增的 $\sigma$ 代数. $X_t:\Omega\mapsto E, X_t\in \mathscr{F}/\mathscr{E}(\forall t\in \boldsymbol{T})$，其中 $\boldsymbol{T}, (\Omega,\mathscr{F},P)$ 和 $(E,\mathscr{E})$ 如定义 1.2 中所设.

(1) 若 $\{X_t, t\in \boldsymbol{T}\}$ 关于 $\{\mathscr{F}_t\}$ 循序可测，则 $\{X_t, t\in \boldsymbol{T}\}$ 是可测的.

(2) 若 $\{E,\rho,\mathscr{E}\}$ 是可测距离空间，即 $\rho$ 是 $E$ 上的一个距离，$\mathscr{E}$ 是 $E$ 上的 Borel $\sigma$ 代数，且对任何 $\omega\in\Omega, X(\bullet,\omega)$ 是右连续的，则 $\{X_t, t\in \boldsymbol{T}\}$ 关于 $\{\mathscr{F}_t\}$ 循序可测.

证明甚易，读者可作为习题验证之.

**定义 1.3** 设 $X = \{X_t, t\in \boldsymbol{T}\}$ 和 $Y = \{Y_t, t\in \boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以 $(E,\mathscr{E})$ 为相空间的两个随机过程. 如果对任何 $t\in \boldsymbol{T}$，都有

$$P(X_t = Y_t) = 1,$$

则称 $X$ 与 $Y$ 是**随机等价**的 (有时简称为等价的)，亦称 $Y$ 是 $X$ 的**修正**. 显然，若 $Y$ 是 $X$ 之修正，则 $X$ 也是 $Y$ 的修正.

虽然这一章研究的对象是离散的时间参数集的随机过程 (这种随机过程有时亦称为随机序列)，但为今后研究连续时间参数集的随机过程，“可分的随机过

程" 的概念, 还是有必要在此作为随机过程的基本概念之一予以介绍, 因为在研究连续时间参数集的随机过程的轨道性质时, 经常要在一个随机等价的随机过程族中选择一个 "较好" 的代表加以研究. 当然好坏的标准很多, 而可分性就是其中重要的一个.

**定义 1.4** 设 $\boldsymbol{T}=[0,\infty)$ 或 $(-\infty,\infty)$ 或 $[a,b], f:\boldsymbol{T}\mapsto\mathbf{R}, D$ 是 $\boldsymbol{T}$ 中一个可数的稠子集. 若对任何 $t\in\boldsymbol{T}$, 都存在 $r_n\in D(n\geqslant 1)$, 使

$$\lim_{n\to\infty} r_n=t,\quad \lim_{n\to\infty} f(r_n)=f(t),$$

则称 $f$ 关于 $D$ 是**可分的**, 而称 $D$ 是 $f$ 的一个可分集. 称 $f$ 是可分的, 如果存在 $f$ 的一个可分集.

显然, $\boldsymbol{T}$ 上的任何右连续函数 $f$ 总是可分的.

设 $X=\{X_t,t\in\boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以 $(\mathbf{R},\mathscr{B}(\mathbf{R}))$ 为相空间的随机过程, $D$ 是 $\boldsymbol{T}$ 的一个可数稠子集. 如果存在 $N\in\mathscr{F}, P(N)=0$, 使得 $\omega\bar{\in}N$ 时, $X(\cdot,\omega)$ 关于 $D$ 是可分的, 则称 $X$ 关于 $D$ 可分, 而称 $N$ 为其例外集, 称 $D$ 为其可分集. 如果存在 $\boldsymbol{T}$ 的一个可数稠子集 $D$, 使 $X$ 关于 $D$ 可分, 则称 $X$ 是**可分的**. 称 $X$ 是完全可分的, 如果 $X$ 关于 $\boldsymbol{T}$ 的任何一个可数稠子集都是可分的.

**定义 1.5** 设 $X=\{X_t,t\in\boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上以可测距离空间 $(E,\rho,\mathscr{E})$ 为相空间的随机过程, 如果对任何 $t_0\in\boldsymbol{T}$, 都有

$$\lim_{\substack{t\to t_0\\ t\in\boldsymbol{T}}} X_t=X_{t_0}, P,$$

则称 $X$ 是**随机连续**的, 如果上述极限中将 $t\to t_0$ 代之以 $t\uparrow t_0$ (或 $t\downarrow t_0$), 则称 $X$ 是**左 (或右) 随机连续**的.

**定理 1.1** 设 $X=\{X_t,t\in\boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以 $(\mathbf{R},\mathscr{B}(\boldsymbol{B}))$ 为相空间的随机过程, $\boldsymbol{T}$ 如定义 1.4 所规定.

(1) $X=\{X_t,t\in\boldsymbol{T}\}$ 总存在一个与之随机等价的可分的随机过程 $Y=\{Y_t,t\in\boldsymbol{T}\}$;

(2) 若 $X$ 是可分的且随机连续的, 则 $X$ 必为完全可分的;

(3) $X$ 是可分的充分必要条件为: 存在 $\boldsymbol{T}$ 的一个可数稠子集 $D$ 和一个 $P$ 零测集 $N, N\in\mathscr{F}, P(N)=0$, 使得对任一闭集 $A\subset\mathbf{R}$ 和开区间 $I\subset\boldsymbol{T}$, 都有

$$\left(\bigcap_{r\in ID} X_r^{-1}(A)-\bigcap_{t\in IT} X_t^{-1}(A)\right)\subset N.$$

**证** 请参见文献 [23] 第 Ⅱ 章定理 2.1~ 定理 2.4.

定理 1.1 的结论 (1) 说明任一随机过程 $X=\{X_t, t\in \boldsymbol{T}\}$ 的可分修正总是存在的; 结论 (2) 说明 $X$ 是完全可分的充分条件是简单的易于检验的; 结论 (3) 说明, 当 $X$ 可分时, 只要概率空间 $(\Omega,\mathscr{F},P)$ 是完备的; 则 $\bigcap\limits_{t\in I\boldsymbol{T}} X_t^{-1}(A)$ 是 $\mathscr{F}$ 可测集, 而对一般的随机过程 $X$, 由于 $I\boldsymbol{T}$ 不是可数集, $\bigcap\limits_{t\in I\boldsymbol{T}} X^{-1}(A)$ 不一定是 $\mathscr{F}$ 可测集, 从而无概率可言. 由定理 1.1 看出: 对于连续时间参数集的随机过程, 在研究其轨道性质时, 可分性是一个很重要的概念, 值得介绍.

关于随机过程的存在性及其构造, 当 $\boldsymbol{T}=[0,\infty)$ 或 $(-\infty,\infty),(E,\mathscr{E})$ 的拓扑结构稍加限制时, 请见 [35] 第四章定理 3.1; 当 $\boldsymbol{T}$ 是可数集, $(E,\mathscr{E})$ 是任意可测空间时, 请见 [35] 第四章定理 3.2. 总之, 本书涉及的随机过程恒存在, 不再赘述.

## §2 Markov 性

人们最初研究的随机过程, 是一种最简单的随机过程 —— 相互独立的随机变量序列 $\{X_n, n=0,1,2,\cdots\}$. 1906 年, 俄罗斯著名数学家 Markov 提出了 “相依” 的随机变量序列的概念, 这种 “相依” 性, 发展成后来的 **“Markov”** 性, 逐渐形成 Markov 过程理论.

在这一章中, 我们研究的随机过程的时间参数集 $\boldsymbol{T}=\{0,1,2,\cdots\}$, 相空间 $(E,\mathscr{E})$ 中的 $E$ 为可数集 (不妨令之为非负整数集), $\mathscr{E}$ 为 $E$ 中全体子集所构成的 $\sigma$ 代数.

**定义 2.1** 称 $E\times E$ 上的矩阵 $P=(p_{i,j}, i,j\in E)$ 为**转移矩阵**, 如果

(1) $p_{i,j}\geqslant 0\ (\forall i,j\in E)$,

(2) $\sum\limits_{j\in E} p_{i,j}=1\ (\forall i\in E)$.

令 $\Omega^*=E^{\boldsymbol{T}}=\mathop{\times}\limits_{i\in \boldsymbol{T}} E_i, E_i\equiv E, \Omega^*$ 中的元素用 $\omega=(\omega(0),\omega(1),\cdots)$ 表之. 记

$$\begin{bmatrix} n \\ A \end{bmatrix}=\{\omega\in\Omega^*:\omega(n)\in A\}\quad (n\in\boldsymbol{T}, A\subset E),$$

特别地, 当 $A=\{j\}$ 是 $E$ 中的单点集时, 记 $\begin{bmatrix} n \\ A \end{bmatrix}=\begin{bmatrix} n \\ j \end{bmatrix}$.

当 $A_k\subset E, t_k\in\boldsymbol{T}(k=1,\cdots,n)$, 则记

$$\begin{aligned}\bigcap_{k=1}^{n}\begin{bmatrix} t_k \\ A_k \end{bmatrix}&=\begin{bmatrix} t_1, & \cdots, t_n \\ A_1, & \cdots, A_n \end{bmatrix}\\ &=\{\omega\in\Omega^*:\omega(t_1)\in A_1,\cdots,\omega(t_n)\in A_n\}.\end{aligned}$$

再令 $\mathscr{F}^* = \sigma\left\{\begin{bmatrix} n \\ j \end{bmatrix} : n \geqslant 0, j \in E\right\}$, 并定义

$$P^i\left(\begin{bmatrix} 0, & 1, & \cdots, & n \\ i_0, & i_1, & \cdots, & i_n \end{bmatrix}\right) = \delta_{i,i_0} p_{i_0,i_1} \cdots p_{i_{n-1},i_n},$$

则由 Kolmogorov 的无穷维乘积空间中造测度的定理 (见 [35] 第四章定理 3.2), $P^i$ 可唯一地扩张到 $\mathscr{F}^*$ 上去而得一概率测度. 对此概率测度, 仍用 $P^i$ 记之. 于是由转移矩阵 $P$ 得到了一族概率空间 $(\Omega^*, \mathscr{F}^*, P^i)(i \in E)$.

再定义一族由 $\Omega^*$ 到 $\Omega^*$ 的推移算子 $\theta_n (n \geqslant 0)$ 如下:

$$(\theta_n \omega)(m) = \omega(n+m) \quad (\forall n \geqslant 0, m \geqslant 0).$$

显然有:

(1) $\theta_n^{-1} \begin{bmatrix} m \\ A \end{bmatrix} = \begin{bmatrix} n+m \\ A \end{bmatrix} (\forall n, m \geqslant 0, A \subset E)$;

(2) $\theta_n^{-1} \mathscr{F}^* \subset \mathscr{F}^* (\forall n \geqslant 0)$.

对任何 $\{t_1, \cdots, t_n\} \subset \boldsymbol{T}, A \in E^n$, 记

$$\begin{matrix} (t_1, \cdots, t_n) \\ A \end{matrix} = \begin{bmatrix} (t_1, \cdots, t_n) \\ A \end{bmatrix} = \{\omega \in \Omega^* : (\omega(t_1), \cdots, \omega(t_n)) \in A\}.$$

有时简记 $P^i(A \bigcap B) = P^i(A, B)$, $\begin{bmatrix} n \\ A \end{bmatrix} = \begin{matrix} n \\ A \end{matrix}$.

**命题 2.1** 对概率空间 $(\Omega^*, \mathscr{F}^*, P^i)$ 有下列 Markov 性:

$$(M): P^i\left(\begin{matrix} (0, 1, \cdots, n-1) \\ A \end{matrix}, \begin{matrix} n \\ j \end{matrix}, \theta_{n+1}^{-1} B\right)$$
$$= P^i\left(\begin{matrix} (0, 1, \cdots, n-1) \\ A \end{matrix}, \begin{matrix} n \\ j \end{matrix}\right) P^j(\theta_1^{-1} B) \tag{2.1}$$

$(\forall n \geqslant 1, i, j \in E, A \in E^n, B \in \mathscr{F}^*)$.

**证** 由于 $E^n$ 是可数集, 故只需证明 (2.1) 对 $A=(i_0,i_1,\cdots,i_{n-1})$ 成立即可. 设 $B=\begin{bmatrix} m \\ k \end{bmatrix}$, 则由 $P^i$ 的定义得:

$$\begin{aligned}
&P^i\left(\begin{matrix}(0,1,\cdots,n-1) \\ A\end{matrix},\begin{matrix} n \\ j\end{matrix},\theta_{n+1}^{-1}\boldsymbol{B}\right)\\
=&P^i\left(\begin{matrix}0,1,\cdots,n-1,n,(n+1+m) \\ i_0,i_1,\cdots,i_{n-1},j,j\end{matrix}\right)\\
=&P^i\left(\begin{matrix}0,\ 1,\ \cdots,\ n-1,\ n \\ i_0,\ i_1,\ \cdots,\ i_{n-1},\ j\end{matrix}\right)P^j\left(\begin{matrix}m+1 \\ k\end{matrix}\right)\\
=&P^i\left(\begin{matrix}(0,1,\cdots,n-1) \\ A\end{matrix},\begin{matrix} n \\ j\end{matrix}\right)P^j\left(\begin{matrix}m+1 \\ k\end{matrix}\right).
\end{aligned}$$

易证使 (2.1) 成立的 $B$ (固定 $n,j,A=(i_0,\cdots,i_{n-1})$ 时) 构成一个 $\sigma$ 代数, 所以对任何 $B\in\mathscr{F}^*$ 成立.

由 $(M)$ 可推出

$$\begin{aligned}
&(M_1):P^i\left(\theta_{n+1}^{-1}B\middle|\begin{matrix}(0,1,\cdots,n-1) \\ A\end{matrix},\begin{matrix} n \\ j\end{matrix}\right)\\
=&P^j(\theta_1^{-1}B) \qquad (2.2)
\end{aligned}$$

$\left(\forall i,j\in E,n\geqslant 1,A\subset E^n,B\in\mathscr{F}^*,P^i\left(\begin{smallmatrix}(0,\cdots,n-1) \\ A\end{smallmatrix},\begin{smallmatrix} n \\ j\end{smallmatrix}\right)>0\right)$.

(2.2) 式的直观意义是: 知道 “现在”, “过去” 与 “将来” 无关. 这就是通常所谓的 Markov 性.

给定一个转移矩阵 $P$, 就得到一族概率空间 $(\Omega^*,\mathscr{F}^*,P^i),(i\in E)$. 再令 $X_n$ 是 $\Omega^*$ 上的坐标函数, 即 $X_n(\omega)=\omega(n)$ $(\forall\omega\in\Omega^*,n\geqslant 0)$, 则称 $(\Omega^*,\mathscr{F}^*,X_n,\theta_m,P^i;i\in E,n,m\geqslant 0)$ 为 $P$ 链, 因为它们是由 $P$ 派生的,$(\Omega^*,\mathscr{F}^*,P^i)(i\in E)$ 称为 $P$ 链的概率空间.

从上面的讨论看出: $\{X_n,n\geqslant 0\}$ 是概率空间 $(\Omega^*,\mathscr{F}^*,P^i)$ 上的以 $(E,\mathscr{E})$ 为状态空间的具有 Markov 性 (2.2) 的随机过程, 也就是说, $\{X_n,n\geqslant 0\}$ 是通常所谓的离散时间的可数状态的 Markov 过程, 即可数状态的 Markov 链. 下面我们再给一般的可数状态的 Markov 链下一个定义.

**定义 2.2** 设 $(\Omega,\mathscr{F},P)$ 为任一概率空间, $E$ 为可数集, $\boldsymbol{T}=\{0,1,2,\cdots\}$ 为时间参数集, $\{X_n,n\in\boldsymbol{T}\}$ 是一族定义在 $(\Omega,\mathscr{F},P)$ 上的取值于 $E$ 的随机变量. 如果对任何 $n\geqslant 1,\{t_1,\cdots,t_n\}\subset\boldsymbol{T},\{i_1,\cdots,i_n\}\subset E$, 且 $\{t_k\}$ 严格上升, 都有

$$\begin{aligned}
&P\left(X_{t_n}=i_n\middle|X_{t_1}=i_1,\cdots,X_{t_{n-1}}=i_{n-1}\right)\\
=&P\left(X_{t_n}=i_n\middle|X_{t_{n-1}}=i_{n-1}\right) \qquad (2.3)
\end{aligned}$$

(当 (2.3) 左边有意义时), 则称 $\{X_n, n \in \boldsymbol{T}\}$ 为一个离散时间可数状态的 Markov 过程, 简称可数状态的 Markov 链. 称 $E$ 为其状态空间. 如果存在一个转移矩阵 $P = (p_{i,j}, i, j \in E)$ 使

$$P(X_n = j \Big| X_{n-1} = i) = p_{i,j} \quad (n \geqslant 1, i, j \in E),$$

(当上式左边有意义时), 则称 $\{X_n, n \in \boldsymbol{T}\}$ 是一个时齐的具有转移矩阵 $P$ 的可数状态的 Markov 链, 简称**时齐的可数状态的 Markov 链**. 称 $P$ 为其**转移概率矩阵** (有时也简称转移矩阵), 称 $\{P(X_0 = i), i \in E\}$ 为其初始分布. 本章所研究的, 都是时齐的可数状态的 Markov 链.

**命题 2.2**　设 $\{X_n, n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一列取值于可数集的随机变量, 则下列三条件等价:

(1) (2.3) 式成立;

(2)
$$\begin{aligned} &P(X_n = i_n | X_0 = i_0, X_1 = i_1, \cdots, X_{n-1} = i_{n-1}) \\ =\ &P(X_n = i_n | X_{n-1} = i_{n-1}) \end{aligned} \tag{2.4}$$

$(\forall n \geqslant 1, \{i_0, i_1, \cdots, i_n\} \subset E)$;

(3)
$$\begin{aligned} &P(X_{t_v} = i_v, n \leqslant v \leqslant m | X_{t_u} = i_u, 1 \leqslant u < n) \\ =\ &P(X_{t_v} = i_v, n \leqslant v \leqslant m | X_{t_{n-1}} = i_{n-1}) \end{aligned} \tag{2.5}$$

$(\forall m \geqslant n \geqslant 1, \{i_1, \cdots, i_m\} \subset E, t_1 < \cdots < t_n < \cdots < t_m)$.

证明甚易, 读者可作为习题验证之.

**定义 2.3**　称两个时齐的可数状态的 Markov 链是等价的, 如果它们具有相同的初始分布和相同的转移概率矩阵.

仅与初始分布和转移概率矩阵有关的性质称为分析性质.　如果只研究 Markov 链的分析性质, 从概率空间 $(\Omega^*, \mathscr{F}^*, P^i)$ 出发即可.

本章所研究的 Markov 链, 都是时齐的可数状态的 Markov 链.

**附注 2.1**　给定 $P$ 链 $(\Omega^*, \mathscr{F}^*, X_n, \theta_m, P^i; i \in E, n, m \geqslant 0)$, 则 $\{X_n, n \geqslant 0\}$ 是概率空间 $(\Omega^*, \mathscr{F}^*, P^i)$ 上的以 $P$ 为转移概率矩阵, 以 $\{\delta_{i,j}, j \in E\}$ 为初始分布的时齐的可数状态的 Markov 链.

## §3　Markov 链的特征数及其性质

令 $\boldsymbol{T} = \{0, 1, 2, \cdots\}$, $E$ 为可数集, $P = (p_{i,j}, i, j \in E)$ 是 $E \times E$ 上的转移矩阵, $(\Omega^*, \mathscr{F}^*, P^i)(i \in E)$ 是 §2 中所定义的 $P$ 链的概率空间.

1. $n$ 步转移概率 $p_{i,j}^{(n)}(n \geqslant 0, i, j \in E)$

令

$$p_{i,j}^{(n)} = P^i\left(\begin{bmatrix} n \\ j \end{bmatrix}\right) \quad (n \geqslant 0, j, i \in E).$$

易证:

$$p_{i,j}^{(n)} = \sum_{\{i_1,\cdots,i_{n-1}\}\subset E} p_{i,i_1}p_{i_1,i_2}\cdots p_{i_{n-1},j} \quad (n \geqslant 2)$$

$$p_{i,j}^{(1)} = P^i\left(\begin{bmatrix} 1 \\ j \end{bmatrix}\right) = p_{i,j};$$

$$p_{i,j}^{(0)} = \delta_{i,j},$$

$p_{i,j}^{(n)}$ 的直观概率意义是从状态 $i$ 出发的 Markov 链经过 $n$ 个时刻后进入状态 $j$ 的转移概率.

**2. 由 $i$ 出发经过 $n$ 个时刻初达状态 $j$ 的概率 $f_{i,j}^{(n)}$**

令

$f_{i,j}^{(n)} = P^i\left(\begin{bmatrix} 1, & 2, & \cdots, & n-1, & n \\ j^c, & j^c, & \cdots, & j^c, & j \end{bmatrix}\right)$ $(n \geqslant 1, i, j \in E, j^c = E - \{j\}$ 表 $\{j\}$ 之补集.)

显然 $f_{i,j}^{(1)} = p_{i,j}^{(1)} = p_{i,j}$.

**3. 由 $i$ 出发经过有限步达 $j$ 的概率 $f_{i,j}^*$.**

令

$$f_{i,j}^* = P^i\left(\bigcup_{n=1}^{\infty}\begin{bmatrix} n \\ j \end{bmatrix}\right) \quad (i, j \in E).$$

显然,

$$f_{i,j}^* = \sum_{n=1}^{\infty} f_{i,j}^{(n)}.$$

4. 初返时间 $T_j(\omega)(j \in E, \omega \in \Omega^*)$.

令

$$T_j(\omega) = \begin{cases} 1, & 若\ \omega \in \begin{bmatrix} 1 \\ j \end{bmatrix}, \\ n, & 若\ \omega \in \begin{bmatrix} 1, & 2, & \cdots, & n-1, & n \\ j^c, & j^c, & \cdots, & j^c, & j \end{bmatrix}, \quad n \geqslant 1, \\ \infty, & 若\ \omega \in \begin{bmatrix} 1, & 2, & \cdots, & n, & \cdots \\ j^c, & j^c, & \cdots, & j^c, & \cdots \end{bmatrix}, \end{cases}$$

则 $T_j$ 是定义在 $(\Omega^*, \mathscr{F}^*, P^i)$ 上的取值于 $\{1, 2, \cdots, \infty\}$ 的随机元, 其概率分布为:

$$P^i(T_j = n) = f_{i,j}^{(n)} \quad (n = 1, 2, \cdots),$$
$$P^i(T = \infty) = 1 - f_{i,j}^*.$$

5. **平均再现时间** $m_{i,i}(i \in E)$.

令

$$m_{i,i} = \begin{cases} \sum_{n=1}^{\infty} n f_{i,i}^{(n)}, & \text{当 } f_{i,i}^* = 1, \\ \infty, & \text{当 } f_{i,i}^* < 1. \end{cases}$$

$m_{i,i}$ 的直观概率意义是: 由状态 $i$ 出发, 第一次返回状态 $i$ 的平均时间. 故称平均再现时间. 注意: 即使 $f_{i,i}^* = 1, m_{i,i}$ 也有可能为 $\infty$.

6. 由 $i$ 出发无穷多次返回 $j$ 的概率 $g_{i,j}(i, j \in E)$.

令

$$g_{i,j} = P^i\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \begin{bmatrix} m \\ j \end{bmatrix}\right) \quad (i, j \in E).$$

下面我们研究这些特征数之间的关系.

**命题 3.1** $g_{i,j} = f_{i,j}^* g_{j,j}$ $(i, j \in E)$.

**证** 由 (2.1) 可得

$$\begin{aligned} g_{i,j} &= P^i\left(\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} \begin{bmatrix} m \\ j \end{bmatrix}\right) \\ &= \sum_{k=1}^{\infty} P^i\left(\begin{bmatrix} 1, & 2, & \cdots, & k-1, & k \\ j^c, & j^c, & \cdots, & j^c, & j \end{bmatrix} \bigcap_{n=k+1}^{\infty} \bigcup_{m=n}^{\infty} \begin{bmatrix} m \\ j \end{bmatrix}\right) \\ &= \sum_{k=1}^{\infty} f_{i,j}^{(k)} q_{jj} = f_{i,j}^* g_{j,j}. \end{aligned}$$

**命题 3.2** $p_{i,j}^{(n)} = \sum\limits_{v=1}^{n} f_{i,j}^{(v)} p_{j,j}^{(n-\nu)}$ $(n \geqslant 1)$. 若令

$$P_{i,j}(\lambda) = \sum_{n=0}^{\infty} p_{i,j}^{(n)} \lambda^n \quad (|\lambda| < 1),$$
$$F_{i,j}(\lambda) = \sum_{n=1}^{\infty} f_{i,j}^{(n)} \lambda^n \quad (|\lambda| \leqslant 1)$$

分别为 $\{p_{i,j}^{(n)}, n \geqslant 0\}$ 和 $\{f_{i,j}^{(n)}, n \geqslant 1\}$ 的母函数, 则还有

$$P_{i,j}(\lambda) = \delta_{i,j} + F_{i,j}(\lambda) P_{j,j}(\lambda), \quad (|\lambda| < 1).$$

特别地, 当 $j=i$ 时,

$$P_{i,i}(\lambda)=\frac{1}{1-F_{i,i}(\lambda)}\quad (|\lambda|<1).$$

**证** 由 (2.1) 即得

$$\begin{aligned}p_{i,j}^{(n)}&=P^i\left(\begin{bmatrix}n\\ j\end{bmatrix}\right)=\sum_{v=1}^{n}P^i\left(\begin{bmatrix}1, & 2, & \cdots, & v-1, & v\\ j^{\mathrm{c}}, & j^{\mathrm{c}}, & \cdots, & j^{\mathrm{c}}, & j\end{bmatrix},\begin{bmatrix}n\\ j\end{bmatrix}\right)\\ &=\sum_{v=1}^{n}P^i\left(\begin{bmatrix}1, & 2, & \cdots, & v-1, & v\\ j^{\mathrm{c}}, & j^{\mathrm{c}}, & \cdots, & j^{\mathrm{c}}, & j^{\mathrm{c}}\end{bmatrix}\right)P^j\left(\begin{bmatrix}n-v\\ j\end{bmatrix}\right)\\ &=\sum_{v=1}^{n}f_{i,j}^{(v)}p_{j,j}^{(n-v)}.\end{aligned}$$

**命题 3.3** $f_{i,j}^{*}=\lim\limits_{\lambda\to 1-0}F_{i,j}(\lambda)=F_{i,j}(1),\ \sum\limits_{n=0}^{\infty}p_{i,j}^{(n)}=\lim\limits_{\lambda\to 1-0}P_{i,j}(\lambda).$

**证** 由定义可直接验证此命题成立.

**命题 3.4** 恒有

(1) $\sum\limits_{n=0}^{\infty}p_{j,j}^{(n)}=\dfrac{1}{1-f_{j,j}^{*}}$ ($\forall j\in E$ 约定 $\dfrac{1}{0}=\infty$);

(2) $m_{j,j}=\begin{cases}\infty, & \text{当 } f_{j,j}^{*}<1,\\ F_{j,j}'(1), & \text{当 } f_{j,j}^{*}=1.\end{cases}$

**证** 由定义可直接验证此命题成立.

**命题 3.5** 若 $q_{i,i}=1, f_{i,j}^{*}>0$, 则 $g_{i,j}=1$.

**证**

$$\begin{aligned}g_{i,j}&=P^i\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\begin{bmatrix}n\\ j\end{bmatrix}\right)\\ &\geqslant P^i\left(\bigcap_{m=1}^{\infty}\bigcup_{n=m}^{\infty}\begin{bmatrix}n\\ j\end{bmatrix},\bigcap_{t=1}^{\infty}\bigcup_{s=t}^{\infty}\begin{bmatrix}s\\ i\end{bmatrix}\right)\\ &=P^i\left(\bigcap_{t=1}^{\infty}\bigcup_{s=t}^{\infty}\begin{bmatrix}s\\ i\end{bmatrix}\right)-P^i\left(\bigcap_{t=1}^{\infty}\bigcup_{s=t}^{\infty}\begin{bmatrix}s\\ i\end{bmatrix},\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\begin{bmatrix}n\\ j^{\mathrm{c}}\end{bmatrix}\right)\\ &=g_{i,i}-\lim_{m\to\infty}\lim_{t\to\infty}P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix}s\\ i\end{bmatrix},\bigcap_{n=m}^{\infty}\begin{bmatrix}n\\ j^{\mathrm{c}}\end{bmatrix}\right).\end{aligned}\tag{3.1}$$

但是, 当 $t>m>0$ 时, 由 (2.1) 有

$$
\begin{aligned}
&P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix} s \\ i \end{bmatrix}, \bigcap_{n=m}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right) \\
\leqslant & \sum_{s=t}^{\infty} P^i\left(\begin{bmatrix} t, & \cdots, & s-1, & s \\ i^{\mathrm{c}}, & \cdots, & i^{\mathrm{c}}, & i \end{bmatrix}, \bigcap_{n=m}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right) \\
\leqslant & \sum_{s=t}^{\infty} P^i\left(\begin{bmatrix} m, & \cdots, & t-1, & t, & \cdots, & s-1, & s \\ j^{\mathrm{c}}, & \cdots, & j^{\mathrm{c}}, & i^{\mathrm{c}}, & \cdots, & i^{\mathrm{c}}, & i \end{bmatrix}, \bigcap_{n=s+1}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right) \\
= & \sum_{s=t}^{\infty} P^i\left(\begin{bmatrix} m, & \cdots, & t-1, & t, & \cdots, & s-1, & s \\ j^{\mathrm{c}}, & \cdots, & j^{\mathrm{c}}, & i^{\mathrm{c}}, & \cdots, & i^{\mathrm{c}}, & i \end{bmatrix}\right) \cdot \left(1-f_{i,j}^{*}\right) \\
\leqslant & P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix} s \\ i \end{bmatrix}, \bigcap_{n=m}^{t-1}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right)\left(1-f_{i,j}^{*}\right).
\end{aligned}
\tag{3.2}
$$

在上述不等式中先令 $t\to\infty$, 次令 $m\to\infty$, 得

$$
\begin{aligned}
&\lim_{m\to\infty}\lim_{t\to\infty} P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix} s \\ i \end{bmatrix}, \bigcap_{n=m}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right) \\
\leqslant & \lim_{m\to\infty}\lim_{t\to\infty} P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix} s \\ i \end{bmatrix}, \bigcap_{n=m}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right)\left(1-f_{i,j}^{*}\right).
\end{aligned}
\tag{3.3}
$$

由于 $\left(1-f_{i,j}^{*}\right)<1$, 所以

$$
\lim_{m\to\infty}\lim_{t\to\infty} P^i\left(\bigcup_{s=t}^{\infty}\begin{bmatrix} s \\ i \end{bmatrix}, \bigcap_{n=m}^{\infty}\begin{bmatrix} n \\ j^{\mathrm{c}} \end{bmatrix}\right)=0. \tag{3.4}
$$

以 (3.4) 代入 (3.1) 得

$$
g_{i,j}=g_{i,i}=1.
$$

**引理 3.1**　设 $\{a_n, n\geqslant 0\}$ 为非负实数序列, 且 $\sum\limits_{n\geqslant 0} a_n>0$, 再设实数序列 $\{b_n, n\geqslant 0\}$ 有极限 $\lim\limits_{n\to\infty} b_n=b$ ($b$ 可以是 $\infty$ 或 $-\infty$). 若

$$
\lim_{n\to\infty}\frac{a_n}{\sum\limits_{\nu=0}^{n} a_\nu}=0,
$$

则

$$
\lim_{n\to\infty}\frac{\sum\limits_{\nu=0}^{n} a_\nu b_{n-\nu}}{\sum\limits_{\nu=0}^{n} a_\nu}=\lim_{n\to\infty} b_n. \tag{3.5}
$$

证明甚易, 读者可作为习题验证之.

**附注 3.1** 若 $\sum\limits_{n=0}^{\infty} a_n < \infty$; 或 $\sum\limits_{n=0}^{\infty} a_n = \infty$, 但 $\{a_n\}$ 有界, 则引理 3.1 中的关于 $\{a_n\}$ 的条件成立.

**命题 3.6** 对任何 $i, j \in E$, 恒有:

$$\lim_{n\to\infty}\left(\sum_{k=1}^{n} p_{i,j}^{(k)} \Big/ \sum_{k=0}^{n} p_{j,j}^{(k)}\right) = f_{i,j}^{*}. \tag{3.6}$$

**证** 由命题 3.2 有

$$\begin{aligned}\sum_{k=1}^{n} p_{i,j}^{(k)} &= \sum_{k=1}^{n}\sum_{\nu=1}^{k} f_{i,j}^{(\nu)} p_{j,j}^{(k-\nu)} \\ &= \sum_{\nu=0}^{n-1} p_{j,j}^{(\nu)} \sum_{k=1}^{n-\nu} f_{i,j}^{(k)}.\end{aligned}$$

取 $a_\nu = p_{j,j}^{(\nu)} (\nu \geqslant 0), b_0 = 0, b_\nu = \sum\limits_{k=1}^{\nu} f_{i,j}^{(k)}, (\nu \geqslant 1)$, 利用引理 3.1 即得命题 3.6.

**命题 3.7** 对任何 $i, j \in E$, 恒有

$$g_{i,j} = f_{i,j}^{*}\left(\lim_{n\to\infty}(f_{j,j}^{*})^{n}\right), \tag{3.7}$$

从而 $g_{i,j}$ 非 0 即等于 $f_{i,j}^{*}$ 且 "$g_{i,i} = 1 \Leftrightarrow f_{i,i}^{*} = 1$".

**证**

$$g_{i,j}(m) = P^{i}\left(\bigcup_{1\leqslant n_1<\cdots<n_m}\bigcap_{s=1}^{m}\begin{bmatrix} n_s \\ j \end{bmatrix}\right) \quad (m \geqslant 1).$$

则由 (2.1) 有

$$\begin{aligned}g_{i,j}(m+1) &= \sum_{k=1}^{\infty} P^{i}\left(\begin{bmatrix} 1, & 2, & \cdots, & k-1, & k \\ j^{c}, & j^{c}, & \cdots, & j^{c}, & j \end{bmatrix}, \bigcup_{k+1\leqslant n_1<\cdots<n_m}\bigcap_{s=1}^{m}\begin{bmatrix} n_s \\ j \end{bmatrix}\right) \\ &= \sum_{k=1}^{\infty} f_{i,j}^{(k)} g_{j,j}(m) \\ &= f_{i,j}^{*} g_{j,j}(m).\end{aligned}$$

若注意: $g_{i,j}(1) = f_{i,j}^{*}, g_{i,j} = \lim\limits_{m\to\infty} g_{i,j}(m)$, 则对 $m$ 作归纳法可证

$$g_{i,j}(m+1) = f_{i,j}^{*}(f_{j,j}^{*})^{m},$$

所以

$$g_{i,j} = \lim_{m\to\infty} g_{i,j}(m+1) = f_{i,j}^{*}\left(\lim_{m\to\infty}(f_{j,j}^{*})^{m}\right).$$

命题 3.7 证毕.

**命题 3.8** 对任何 $i,j\in E$, 恒有

$$f^*_{i,j}=p_{i,j}+\sum_{k\neq j}p_{i,k}f^*_{k,j}. \tag{3.8}$$

**证** 由 (2.1) 有

$$\begin{aligned}f^*_{i,j}&=P^i\left(\bigcup_{n=1}^{\infty}\begin{bmatrix}n\\ j\end{bmatrix}\right)\\&=p_{i,j}+\sum_{n=2}^{\infty}P^i\left(\begin{bmatrix}1, & 2, & \cdots, & n-1, & n\\ j^c, & j^c, & \cdots, & j^c, & j\end{bmatrix}\right)\\&=p_{i,j}+\sum_{n=2}^{\infty}\sum_{k\neq j}P^i\left(\begin{bmatrix}1\\ k\end{bmatrix}\right)P^k\left(\begin{bmatrix}1, & \cdots, & n-2, & n-1\\ j^c, & \cdots, & j^c, & j\end{bmatrix}\right)\\&=p_{i,j}+\sum_{k\neq j}p_{i,k}f^*_{k,j}.\end{aligned}$$

## §4 状态的分类及判别准则

在这一节中, 恒设 $(\Omega^*,\mathscr{F}^*,X_n,\theta_m,P^i,i\in E,m,n\geqslant 0)$ 是 $P$ 链, $P=(p_{i,j},i,j\in E)$ 是 $P$ 链的转移概率矩阵,$E$ 是其状态空间, 其中每个点 $i$ 都称为一个状态, $\{\theta_m,m\geqslant 0\}$ 是 $P$ 链的推移算子族, $(\Omega^*,\mathscr{F}^*,P^i)(i\in E)$ 是 $P$ 链的状态空间. 我们将进行状态分类, 并给出各类状态的判别准则. 记 $P^{(n)}=P^n=(p^{(n)}_{i,j},i,j\in E)$.$f^{(n)}_{i,j},f^*_{i,j},g_{i,j},m_{i,i},T_i$ 和 $F_{i,j}(\lambda)$ 的定义见 §3.

**定义 4.1** 称状态 $i$ **可达状态** $j$ (记之为 $i\rightsquigarrow j$), 如果存在正整数 $m$, 使 $p^{(m)}_{i,j}>0$. 若 $i\rightsquigarrow j$ 且 $j\rightsquigarrow i$, 则称 $i,j$ 互逆, 记之为 $i\leftrightsquigarrow j$. 如果 $f^*_{i,i}<1$; 则称 $i$ 是**暂留状态**; 如果 $f^*_{i,i}=1$, 则称 $i$ 是**常返状态**. 记全体暂留状态为 $N$, 记全体常返状态为 $R$. 若常返状态 $i$ 满足 $m_{i,i}<\infty$, 则称 $i$ 是**正常返状态**, 反之称 $i$ 为**零常返状态**, 简称零状态. 记全体正常返状态为 $R^+$; 全体零常返状态为 $R^0$.

**命题 4.1** 对任何状态 $i,j$, 总有:

$$i\rightsquigarrow j\Leftrightarrow f^*_{i,j}>0;\quad i\leftrightsquigarrow j\Leftrightarrow f^*_{i,j}f^*_{j,i}>0.$$

**证** 由

$$\sup_{n\geqslant 1}p^{(n)}_{i,j}\leqslant f^*_{i,j}\leqslant\sum_{n=1}^{\infty}p^{(n)}_{i,j}$$

立即可证出此命题.

**命题 4.2** $i\in N\Leftrightarrow\sum\limits_{n=0}^{\infty}p^{(n)}_{i,i}<\infty.$

**证** 由命题 3.4 立得命题 4.2.

**命题 4.3** 设 $i \neq j, f^*_{i,i} = 1, f^*_{i,j} > 0$, 则

(1) 存在正整数 $m$ 及正数 $c$, 使

$$p^{(m+n)}_{j,j} \geqslant c p^{(n)}_{i,i} \quad (\forall n \geqslant 0);$$

(2) $f^*_{i,i} = f^*_{j,j} = f^*_{i,j} = f^*_{j,i} = 1$;

(3) $\sum\limits_{n=1}^{\infty} p^{(n)}_{i,i} = \sum\limits_{n=1}^{\infty} p^{(n)}_{j,j} = \sum\limits_{n=1}^{\infty} p^{(n)}_{i,j} = \sum\limits_{n=1}^{\infty} p^{(n)}_{j,i} = \infty$.

**证** (1) 因 $f^*_{i,j} > 0$, 故存在 $M \geqslant 1$, 使 $\alpha = p^{(M)}_{i,j} > 0$. 用 Markov 性 (2.1) 及命题 3.7 有

$$\begin{aligned}
\alpha &= P^i\left(\begin{bmatrix} M \\ j \end{bmatrix}\right) \\
&= P^i\left(\begin{bmatrix} M \\ j \end{bmatrix}, \bigcup_{n>M} \begin{bmatrix} n \\ i \end{bmatrix}\right) + P^i\left(\begin{bmatrix} M, & M+1, & M+2, & \cdots \\ j, & i^c, & i^c, & \cdots \end{bmatrix}\right) \\
&= P^i\left(\begin{bmatrix} M \\ j \end{bmatrix}\right) P^j\left(\bigcup_{n=1}^{\infty} \begin{bmatrix} n \\ i \end{bmatrix}\right) + P^i\left(\begin{bmatrix} M, & M+1, & M+2, & \cdots \\ j, & i^c, & i^c, & \cdots \end{bmatrix}\right) \\
&\leqslant \alpha f^*_{j,i} + P^i\left(\bigcap_{n=M+1}^{\infty} \begin{bmatrix} n \\ i^c \end{bmatrix}\right) \\
&\leqslant \alpha f^*_{j,i} + P^i\left(\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \begin{bmatrix} n \\ i^c \end{bmatrix}\right) \\
&= \alpha f^*_{j,i} + (1 - g_{i,i}) = \alpha f^*_{j,i}.
\end{aligned}$$

由 $\alpha > 0$ 和 $0 \leqslant f^*_{j,i} \leqslant 1$ 得 $f^*_{j,i} = 1$. 所以存在 $M'$, 使 $\beta = p^{(M')}_{j,i} > 0$. 于是

$$p^{(M+n+M')}_{j,j} \geqslant p^{(M')}_{j,i} p^{(n)}_{i,i} p^{(M)}_{i,j} = \alpha\beta p^{(n)}_{i,i}.$$

所以, 若取 $m = M + M', c = \alpha\beta$ 即满足 (1) 的要求. (1) 得证.

(2) 因为 $f^*_{i,i} = 1$, 即 $i \in R$. 由命题 4.2 知 $\sum\limits_{n=0}^{\infty} p^{(n)}_{i,i} = \infty$. 再用本命题中的 (1) 得 $\sum\limits_{n=0}^{\infty} p^{(n)}_{j,j} = \infty$, 从而 $j \in R$, 故 $f^*_{j,j} = 1$. 总之, 我们由 $f^*_{i,i} = 1$ 和 $f^*_{i,j} > 0$ 推出了: $f^*_{j,j} = f^*_{j,i} = 1$. 类似地, 由 $f^*_{j,j} = 1$ 和 $f^*_{j,i} = 1$ 亦可推出 $f^*_{i,j} = 1$. (2) 证毕.

由命题 4.2 及命题 3.6 可证 (3) 亦成立. 命题 4.3 证毕.

**定义 4.2** 设 $Q = (q_{i,j}, i, j \in E)$ 是 $E \times E$ 上之矩阵, $E_1 \subset E$, 称 $Q_1 = (q_{i,j}, i, j \in E_1)$ 为 $Q$ 在 $E_1$ 上的局限, 记作 $Q_1 = Q$ on $E_1$. (对于行, 列向量, 亦有类似定义.)

**定理 4.1** 对于任何 $P$ 链而言, 其状态空间 $E$ 与转移概率矩阵 $P$ 恒有下述形式:

$$E = N\bigcup R = N\bigcup(R^0\bigcup R^+);$$

$$P = \begin{array}{c} \\ N \\ R^0 \\ R^+ \end{array} \begin{array}{c} \begin{array}{ccc} N & R^0 & R^+ \end{array} \\ \begin{pmatrix} Q & Q_0 & Q_+ \\ 0 & P_0 & 0 \\ 0 & 0 & P_+ \end{pmatrix} \end{array},$$

即 $P$ on $R^0 = P_0, P$ on $R^+ = P_+, P$ on $N = Q, Q_0 = (p_{i,j}, i \in N, j \in R^0), Q_+ = (p_{i,j}, i \in N, j \in R^+), p_{i,j} = 0$ (当 $i \in R^0, j \overline{\in} R^0), p_{i,j} = 0$ (当 $i \in R^+, j\overline{\in} R^+), P_0$ 和 $P_+$ 分别为定义在 $R^0 \times R^0$ 和 $R^+ \times R^+$ 上的转移矩阵.

**证** (1) 若 $i \in R$, 即 $f^*_{i,i} = 1$. 若 $f^*_{i,j} > 0$, 则由命题 4.3 有 $f^*_{j,j} = 1$, 即 $j \in R$. 这就证明了:

$$i \in R \text{ 且 } j \in N \Rightarrow f^*_{i,j} = 0 \Rightarrow p_{i,j} = 0.$$

(2) 若 $i \in R^+$ (从而 $f^*_{i,i} = 1, m_{i,i} < \infty$), 且 $f^*_{i,j} > 0$, 则由命题 4.4 知:

$$f^*_{j,i} = f^*_{j,j} = f^*_{j,i} = 1,$$

且存在正整数 $m$ 和正数 $c$, 使

$$p^{(m+n)}_{j,j} \geqslant c p^{(n)}_{i,i} \quad (\forall n \geqslant 0).$$

再用命题 3.2 及命题 3.4 有

$$\begin{aligned} \frac{1}{m_{j,j}} &= \lim_{\lambda \uparrow 1}(1-\lambda)P_{j,j}(\lambda) \\ &= \lim_{\lambda \uparrow 1}(1-\lambda)\sum_{n=0}^{\infty} p^{(n)}_{j,j}\lambda^n \\ &\geqslant \lim_{\lambda \uparrow 1} c(1-\lambda)\sum_{n=0}^{\infty} p^{(n)}_{i,i}\lambda^{n+m} \\ &= \frac{c}{m_{i,i}} > 0. \end{aligned}$$

所以 $j \in R^+$. 这就证明了:

$$i \in R^+ \text{ 和 } j\overline{\in} R^+ \Rightarrow f^*_{i,j} = 0 \Rightarrow p_{i,j} = 0.$$

(3) 若 $i \in R^0, f^*_{i,j} > 0$, 则由命题 4.3 (1) 有 $f^*_{j,j} = 1$. 谬设 $j \in R^+$, 则由 (2) 知 $i \in R^+$. 这与假设矛盾. 于是我们证明了:

$$i \in R^0 \text{ 和 } j \in R^+ \Rightarrow f^*_{i,j} = 0 \Rightarrow p_{i,j} = 0.$$

综合上述三步, 定理 4.1 得证.

下面我们对 $R^0$ ($R^+$ 亦类似) 再分类. 如定义 4.1, 对任何 $i,j\in R^0$, "$i\leftrightsquigarrow j\Leftrightarrow f^*_{i,j}f^*_{j,i}>0$". 显然, 关系 "$\leftrightsquigarrow$" 定义在 $R^0$ 上成一等价关系, 即是 (i) 具有自返性 ($i\in R^0\Rightarrow i\leftrightsquigarrow i$); (ii) 具有对称性 ($i,j\in R^0, i\leftrightsquigarrow j\Rightarrow j\leftrightsquigarrow i$); (iii) 具有传递性 ($i,j,k\in R^0, i\leftrightsquigarrow j, j\leftrightsquigarrow k\Rightarrow i\leftrightsquigarrow k$). 所以按关系 $\leftrightsquigarrow$ 可以把 $R^0$ 分成一些互不相交的等价类:

$$R^0=R^0_1\bigcup R^0_2\bigcup\cdots.$$

由命题 4.3 得知: 若 $i,j\in R^0_n$, 则 $f^*_{i,j}=1$; 若 $i\in R^0_n, j\in R^0_m, n\neq m$, 则 $f^*_{i,j}=0$, 所以我们有

**定理 4.2** 对状态空间 $E$, 总可分解成:

$$\begin{aligned}E&=N\bigcup R=N\bigcup(R^0\bigcup R^+)\\&=N\bigcup\left(\bigcup_m R^0_m\right)\bigcup\left(\bigcup_n R^+_n\right),\end{aligned}$$

而且转移矩阵 $P$ 总具有下述形状:

$$\begin{array}{c}\\ N\\ R^0_1\\ R^0_2\\ \vdots\\ R^+_1\\ R^+_2\\ \vdots\end{array}
\begin{array}{c}\begin{array}{ccccccc}N & R^0_1 & R^0_2 & & R^+_1 & R^+_2 & \cdots\end{array}\\
\left(\begin{array}{ccccccc}
Q & Q_{0,1} & Q_{0,2} & \cdots & Q_{+,1} & Q_{+,2} & \cdots\\
0 & P_{0,1} & 0 & & 0 & 0 & \cdots\\
0 & 0 & P_{0,2} & \cdots & 0 & 0 & \cdots\\
\vdots & \vdots & \vdots & & \vdots & \vdots & \cdots\\
0 & 0 & 0 & & P_{+,1} & 0 & \cdots\\
0 & 0 & 0 & & 0 & P_{+,2} & \cdots\\
\vdots & \vdots & \vdots & & \vdots & \vdots &
\end{array}\right)\end{array}$$

此即 $P$ on $R^0_m=P_{0,m}, P$ on $R^+_n=P_{+,n}, P$ on $N=Q, Q_{0,m}=(p_{i,j}, i\in N, j\in R^0_m), Q_{+,n}=(p_{i,j}, i\in N, j\in R^+_n), p_{i,j}=0$ (当 $i\in R^0_m, j\overline{\in}R^0_m$), $p_{i,j}=0$ (当 $i\in R^+_n, j\in R^+_n$), $P_{0,m}$ 和 $P_{+,n}$ 分别为 $R^0_m\times R^0_m$ 和 $R^+_n\times R^+_n$ 上的转移矩阵, 每一个 $R^0_m(R^+_n)$ 都称为一个 "零类" ("正类"). 零类和正类统称为常返类. $N$ 中每一个状态自成一个类.

此定理的一切结论都是显然的.

**定义 4.3** 称状态空间 $E$ 中的非空子集 $J$ 是**封闭的**, 如果 "$i\in J, j\overline{\in}J\Rightarrow p_{i,j}=0$" (即是, "$i\in J\Rightarrow\sum\limits_{j\in J}p_{i,j}=1$").

**定义 4.4** 对任何 $J \subset E$, 记 $\widehat{J} = \{i \in E : f^*_{i,J} = 0\}$, 其中 $f^*_{i,J} = P^i\left(\bigcup\limits_{n=1}^{\infty}\begin{bmatrix} n \\ J \end{bmatrix}\right)$, 其概率意义为: 从状态 $i$ 出发, 经过有限步达状态集 $J$ 的概率, 而 $\widehat{J}$ 的概率意义是: 不能经过有限步到达状态集 $J$ 的状态的全体.

**命题 4.4** 设 $J$ 和 $\widehat{J}$ 都是 $E$ 的非空子集, 则

(1) $\widehat{J}$ 是封闭的;

(2) $J^c \bigcap \widehat{J}$ 也是封闭的.

**证** (1) 设 $i \in \widehat{J}, j \overline{\in} \widehat{J}$, 则

$$p_{i,j} f^*_{j,J} \leqslant f^*_{i,J} = 0, \quad f^*_{j,J} > 0.$$

所以 $p_{i,j} = 0$, 此即 $\widehat{J}$ 是封闭的.

(2) 取 $i \in \widehat{J}$. (a) 若 $j \overline{\in} \widehat{J}$, 则由 (1) 可知 $p_{i,j} = 0$. (b) 若 $j \in J$, 则 $p_{i,j} \leqslant f^*_{i,J} = 0$. 由 (a)、(b) 可知 $\sum\limits_{j \in \widehat{J} \cap J^c} p_{i,j} = 1$, 此即 $\widehat{J} \bigcap J^c$ 是封闭的. 命题 4.4 证毕.

**命题 4.5** 对任何 $i \in E, \{j \in E : f^*_{i,j} > 0\}$ 是封闭的.

**证** 取 $k$ 和 $l$ 使 $f^*_{i,k} > 0$ 而 $f^*_{i,l} = 0$, 则 $f^*_{k,l} = 0$, 从而 $p_{k,l} = 0$, 此即 $\{j \in E : f^*_{i,j} > 0\}$ 是封闭的.

**定义 4.5** 称状态空间 $E$ (或者称 $P$) 是**可约的**, 如果存在 $E$ 的一个封闭的真子集 $J$, 反之称 $E$ 是**不可约的**.

**命题 4.6** $E$ 不可约的充分必要条件是: 对任何 $i, j \in E$, 都有 $f^*_{i,j} > 0$.

**证** 充分性显然成立. 必要性也只需注意: 若有 $i_0, j_0 \in E$, 使 $f^*_{i_0,j_0} = 0$, 则 $\{j \in E : f^*_{i_0,j} > 0\}$ 是 $E$ 的一个封闭的真子集.

**推论 4.1** 若存在 $i_0 \in E$, 使 $f^*_{i_0,j} f^*_{j,i_0} > 0 (\forall j \in E)$, 则 $E$ 是不可约的.

**定义 4.6** 设 $i \in E, i \rightsquigarrow i$, 称 $\{n \geqslant 1 : p^{(n)}_{i,i} > 0\}$ 的最大公因子

$$\text{G.C.D.} \quad \{n \geqslant 1 : p^{(n)}_{i,i} > 0\}$$

为 $i$ 之**周期**, 记之为 $d_i$.

**命题 4.7** 若 $f^*_{i,j}, f^*_{j,i} > 0$, 则 $d_i = d_j$.

**证** 由假设 $i \rightsquigarrow i, j \rightsquigarrow j, i \rightsquigarrow j$, 所以 $d_i$ 和 $d_j$ 都有定义, 且存在正整数 $s$ 和 $t$ 使

$$p^{(s)}_{i,j} p^{(t)}_{j,i} > 0.$$

于是

$$p_{j,j}^{(n)} > 0 \Rightarrow p_{i,i}^{(s+n+t)} \geqslant p_{i,j}^{(s)} p_{j,j}^{(n)} p_{j,i}^{(t)} > 0.$$

所以 $d_i$ 能整除 $(s+n+t)$. 但由 $p_{j,j}^{(n)} > 0$ 知 $p_{j,j}^{(2n)} > 0$, 所以由 $p_{j,j}^{(n)} > 0$ 也可推出 $d_i$ 能整除 $(s+2n+t)$. 于是有

$$p_{j,j}^{(n)} > 0 \Rightarrow d_i \text{ 能整除 } n,$$

从而

$$d_i \text{ 能整除 } d_j.$$

由 $i$ 和 $j$ 地位的对称性可知 $d_i = d_j$.

**附注 4.1** 命题 4.7 说明: 当 $i \in N, f_{i,i}^* > 0$ 时 $d_i$ 有定义; 当 $i, j \in R_n^+$ (或 $R_m^0$) 时, $d_i$ 和 $d_j$ 都有定义且相等.

下面把 $R_m^0$ (或 $R_n^+$) 分成一些更小的循环子类.

由命题 4.7 知: 任取 $i \in R_m^0, d_i$ 有定义且不依赖 $i \in R_m^0$. 故可定义 $d_i$ 为类 $R_m^0$ 之周期, 用 $d^0(m)$ 表之, 以说明它仅与 $R_m^0$ 有关, 而与 $R_m^0$ 中的具体状态 $i$ 无关. 若 $d^0(m) = 1$, 则称 $R_m^0$ 是无周期的.

**命题 4.8** 设 $f_{i,j}^* f_{j,i}^* > 0$, 则

(1) $p_{i,j}^{(m)} p_{j,i}^{(n)} > 0 \Rightarrow d_i | (m+n)$,

其中 $d|k$ 表示 $d$ 能整除 $k$;

(2) $p_{i,j}^{(n_1)} p_{i,j}^{(n_2)} > 0 \Rightarrow d_i | (n_1 - n_2)$;

(3) 对命题中之 $i$, 存在 $M(i)$ 使

$$p_{i,i}^{(md_i)} > 0 \quad (\text{对一切 } m \geqslant M(i));$$

(4) 对命题中之 $i, j$, 存在 $r_j$ 使

$$p_{i,j}^{(n)} > 0 \Rightarrow n = r_j \quad (\text{mod} d_i);$$

(5) 对命题中之 $i, j$, 存在 $r_j$ 和 $K(i)$, 使

$$p_{i,j}^{(nd_i + r_j)} > 0 \quad (\text{对一切 } n \geqslant K(i)).$$

**证** (1) 因为

$$p_{i,j}^{(m)} p_{j,i}^{(n)} > 0 \Rightarrow p_{i,i}^{(m+n)} > 0,$$

所以 $d_i | (m+n)$.

(2) 由 $f_{j,i}^* > 0$ 知: 存在正整数 $v$ 使 $p_{j,i}^{(v)} > 0$. 所以, 由 $p_{i,j}^{(n_1)} p_{i,j}^{(n_2)} > 0$ 推知

$$p_{i,j}^{(n_1)} p_{j,i}^{(v)} > 0, \quad p_{i,j}^{(n_2)} p_{j,i}^{(v)} > 0.$$

由 (1) 知

$$d_i|(n_1+v), \quad d_i|(n_2+v).$$

所以 $d_i|(n_1-n_2)$.

(3) 由 $d_i$ 之定义知: 存在 $n_1,\cdots,n_k$, 使

$$\prod_{s=1}^{k} p_{i,i}^{(n_s)} > 0, \quad d_i = \text{G.C.D.} \ \{n_1,\cdots,n_k\}.$$

由数论中一条定理可知: 存在 $M(i)$, 当 $m \geqslant M(i)$ 时有

$md_i = \sum_{s=1}^{k} c_s n_s$ ($c_s$ 是正整数). 所以

$$p_{i,i}^{(md_i)} \geqslant \prod_{s=1}^{k} (p_{i,i}^{(n_s)})^{c_s} > 0 \quad (\text{当 } m \geqslant M(i)).$$

(4) 可由 (2) 立即推得.

(5) 由 $f_{i,j}^* > 0$ 及 (4) 得知: 存在 $m'd_i + r_j$, 使

$$p_{i,j}^{(m'd_i+r_j)} > 0.$$

由 (3), 取 $K(i) = m' + M(i)$ 即为所求.

**定义 4.7** 设 $E$ 是不可约的, $d$ 为其周期 (由命题 4.7, $E$ 中任一状态 $i$ 之周期 $d_i$ 不依赖 $i$, 故可以用 $d$ 表示 $E$ (或者 $P$) 之周期.), 若 $d > 1$, 令 $H_{i,j} = \text{G.C.D.}\{n \geqslant 1 : p_{i,j}^{(n)} > 0\}$, 若 $d|H_{i,j}$, 则称 $i,j$ 具有关系 $\approx$, 记作 $i \approx j$. 若 $d = 1$, 则称 $E$ (或者 $P$) 是无周期的.

**命题 4.9** 定义 4.7 中确定的关系 $\approx$ 是一个等价关系.

**证** (1) 自返性. 由 $d = H_{i,i} (\forall i \in E)$ 知 $\approx$ 具有自返性.

(2) 对称性. 设 $i \approx j$. 任取 $m$ 满足 $p_{i,j}^{(m)} > 0$, 必有 $d|m$. 而由命题 4.8 (1) 知: 若 $p_{j,i}^{(n)} > 0$, 则 $d|(m+n)$. 所以 $d|n$, 从而 $d|H_{j,i}$, 即 $j \approx i$. 对称性证毕.

(3) 传递性. 设 $i \approx j, j \approx k$. 任取 $s$ 和 $t$ 使 $p_{i,j}^{(s)} > 0, p_{j,k}^{(t)} > 0$, 必有 $d|s$ 且 $d|t$ 及 $p_{i,k}^{(s+t)} > 0$. 由命题 4.8 (2) 知:

$$p_{i,k}^{(n)} > 0 \Rightarrow n = s + t \pmod{d}.$$

再注意: $d|(s+t)$, 可知

$$p_{i,k}^{(n)} > 0 \Rightarrow n = 0 \pmod{d},$$

此即 $i \approx k$. 命题 4.9 证毕.

**命题 4.10** 设 $E$ 是不可约的, $i \approx j, d$ 为 $E$ 之周期, 则 $d = H_{i,j}$.

**证** 设 $H_{i,j} = \delta$, 任取 $t$ 使 $p_{i,j}^{(t)} > 0$. 由命题 4.8 (3) 可取 $m_0$, 使

$$p_{j,j}^{(m_0 d)} p_{j,j}^{((m_0+1)d)} > 0,$$

于是

$$p_{i,j}^{(t+m_0 d)} p_{i,j}^{(t+(m_0+1)d)} > 0.$$

所以

$$\delta | (t + m_0 d), \text{ 且 } \delta | (t + (m_0+1)d),$$

从而 $\delta | d$. 而 $i \approx j$, 所以 $d | \delta$. 总之 $d = \delta$. 命题 4.10 证毕.

**命题 4.11** 设 $E$ 不可约, $d$ 为其周期, 且 $p_{i,j}^{(m)} > 0, p_{i,k}^{(n)} > 0, m = n (\text{mod } d)$, 则 $j \approx k$.

**证** 由本命题之假设及命题 4.8 (2) 可知:

$$\text{“} p_{j,k}^{(t)} > 0 \Rightarrow p_{i,k}^{(t+m)} > 0 \Rightarrow (t+m) = n (\text{mod } d) \text{”},$$

但假设了 $m = n (\text{mod } d)$, 所以

$$\text{“} p_{j,k}^{(t)} > 0 \Rightarrow t = 0 \quad (\text{mod } d) \text{”},$$

故 $j \approx k$. 命题证毕.

**命题 4.12** 设 $E$ 不可约且其周期为 $d$, 则关系 $\approx$ 恰巧把 $E$ 分成了 $d$ 个**循环子类**:

$$E = C_1 \bigcup C_2 \bigcup \cdots \bigcup C_d,$$

$i, j$ 同属于一个循环类 $C_r$ 的充分必要条件是: $i \approx j$, 而且

$$\text{“} i \in C_r \text{ 且 } j \overline{\in} C_{r+1} \Rightarrow p_{i,j} = 0 \text{”},$$

从而

$$p_{i,j}^{(s)} > 0, i \in C_r \text{ 且 } j \in C_q \Rightarrow q - r = s (\text{mod } d).$$

**证** 不妨设 $d > 1$. 由于 $\approx$ 是一个等价关系, 所以 $\approx$ 可把 $E$ 分成若干个循环类, 使两个状态 $i$ 和 $j$ 同属于一个循环类的充分必要条件是 $i \approx j$. 故为证此命题只需证明两点: (1) 循环类的个数恰为 $d$; (2) $i \in C_r, j \overline{\in} C_{r+1} \Rightarrow p_{i,j} = 0$. (注: 当 $1 \leqslant r \leqslant d, m = r (\text{mod } d)$ 时, 记 $C_m = C_r$.)

(1) 先证循环类之个数不超过 $d$. 反之, 必有状态 $i_0, i_1, \cdots, i_d$, 使 $i_s$ 与 $i_t$ 无关系 $\approx$ (此处 $s \neq t, 0 \leqslant s, t \leqslant d$). 取 $m_1, \cdots, m_d$ 使

$$\prod_{k=1}^{d} p_{i_0,i_k}^{(m_k)} > 0$$

由命题 4.11 知

$$m_s \neq m_t \pmod d$$

对一切 $s \neq t, 1 \leqslant s, t \leqslant d$ 成立.

由命题 4.8 (3), 可再取 $m_0$ 使

$$p_{i_0,i_0}^{(m_0 d)} > 0.$$

由命题 4.11 又有

$$m_0 d \neq m_s \pmod d \quad (1 \leqslant s \leqslant d).$$

而 $d$ 个正整数 $\{m_1, \cdots, m_d\}$ 中的每一个都不能用 $d$ 整除, 且互不同余, 这是不可能的. 这就证明了循环类之个数不能超过 $d$.

于是可令

$$E = C_1 \bigcup C_2 \bigcup \cdots \bigcup C_d.$$

再证每个 $C_r$ 都非空 $(r = 1, 2, \cdots, d)$. 由命题 4.10 知: $i$ 和 $j$ 属于同一个循环类蕴涵了: $p_{i,j} = 0$. 今在各个类中取一个非空者, 认为这是 $C_1$. 任取 $i_1 \in C_1$, 由于

$$1 = \sum_{j \in E} p_{i_1,j} = \sum_{j \in \bigcup\limits_{k=2}^{d} C_k} p_{i_1,j},$$

所以必存在 $i_2 \in \bigcup\limits_{k=2}^{d} C_k$, 使 $p_{i_1,i_2} > 0$. 就认为含 $i_2$ 者是 $C_2$. 若 $d = 2$, 则终止. 若 $d > 2$, 由命题 4.10 知

$$p_{i_1,j}^{(2)} = 0 \quad (\forall j \in C_1).$$

而

$$p_{i_1,i_2} > 0, \quad p_{i_1,j}^{(2)} \geqslant p_{i_1,i_2} p_{i_2,j} \quad (j \in C_1),$$

所以

$$p_{i_2,j} = 0 \quad (\forall j \in C_1).$$

而 $p_{i_2,j} = 0$ $(\forall j \in C_2)$, 所以

$$1 = \sum_{j \in E} p_{i_2,j} = \sum_{j \in \bigcup\limits_{k=3}^{d} C_k} p_{i_2,j}.$$

因此，存在 $i_3 \in \bigcup_{k=3}^{d} C_k$，使 $p_{i_2,i_3} > 0$. 就认为含 $i_3$ 者为 $C_3$. 若 $d = 3$，则终止. 若 $d > 3$，则仿上继续进行下去. $d$ 次以后，得到 $d$ 个状态 $i_1, \cdots, i_d$ 满足 $i_k \in C_k$ $(k = 1, \cdots, d)$，且

$$\prod_{k=1}^{d-1} p_{i_k,i_{k+1}} > 0.$$

(1) 证毕.

(2) 再证 $i \in C_r, j \overline{\in} C_{r+1} \Rightarrow p_{i,j} = 0$.

设 $i \in C_\alpha, j \in C_\beta, p_{i,j} > 0$. 往证 $\beta = (\alpha+1) \pmod d$. 事实上，不妨令 $\beta \geqslant \alpha$. 因为 $i \approx i_\alpha, j \approx i_\beta$.($i_1, \cdots, i_d$ 是 (1) 中找出的那组状态，故有 $p_{i_k,i_{k+1}} > 0$.) 所以存在 $m$ 和 $m'$ 使

$$p_{i,i_\alpha}^{(md)} > 0, \quad p_{i_\beta,j}^{(m'd)} > 0.$$

因此

$$p_{i,j}^{(md+\beta-\alpha+m'd)} \geqslant p_{i,i_\alpha}^{(md)} p_{i_\alpha,i_{\alpha+1}} \cdots p_{i_{\beta-1},i_\beta} p_{i_\beta,j}^{(m'd)} > 0.$$

但是又有

$$p_{i,j} > 0.$$

所以由命题 4.8 (2) 知

$$md + \beta - \alpha + m'd = 1 \pmod d,$$

从而

$$\beta - \alpha = 1 \pmod d.$$

命题证毕.

**定理 4.3** 设 $E$ 不可约，其周期为 $d$，则 $E$ 可分成 $d$ 个互不相交的非空的循环子类：$E = C_1 \bigcup C_2 \bigcup \cdots \bigcup C_d$，这时 $P$ 有如下形式 (自然设 $d > 1$)：

$$\begin{array}{c} \\ C_1 \\ C_2 \\ \vdots \\ C_{d-1} \\ C_d \end{array}
\begin{array}{c} \begin{array}{cccccc} C_1 & C_2 & C_3 & & C_{d-1} & C_d \end{array} \\
\begin{pmatrix} 0 & \widehat{P}_1 & 0 & & 0 & 0 \\ 0 & 0 & \widehat{P}_2 & & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & 0 & \widehat{P}_{d-1} \\ \widehat{P}_d & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \end{array}$$

$P^{(d)} \overset{\text{def.}}{=} P^d \overset{\text{def.}}{=} \widetilde{P} \overset{\text{def.}}{=} (\widetilde{p}_{i,j}, i, j \in E)$ 有如下形式:

$$\begin{array}{c} \\ C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_d \end{array} \begin{array}{c} \begin{array}{ccccc} C_1 & C_2 & C_3 & & C_d \end{array} \\ \begin{pmatrix} \widetilde{P}_1 & 0 & 0 & & 0 \\ 0 & \widetilde{P}_2 & 0 & & 0 \\ 0 & 0 & \widetilde{P}_3 & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \widetilde{P}_d \end{pmatrix} \end{array}$$

即是 $\widetilde{P}$ on $C_\alpha = \widetilde{P}_\alpha$ 是转移矩阵, $\alpha = 1, 2, \cdots, d, \widetilde{p}_{i,j} = 0$ (当 $i \in C_\alpha, j \overline{\in} C_\alpha, \alpha = 1, 2, \cdots, d$). 此外,$\widetilde{P}_\alpha$ 是不可约的、无周期的.

**证** 只需证 $\widetilde{P}$ on $C_\alpha = \widetilde{P}_\alpha$ 是不可约的无周期的转移矩阵. (其他论断前面诸命题已经证明了.)

任取 $i \in C_\alpha$. 若 $\widetilde{p}_{i,j} > 0$, 即 $p_{i,j}^{(d)} > 0$, 则必存在 $j_1, \cdots, j_{d-1}$, 使

$$p_{i,j_1} p_{j_1,j_2} \cdots p_{j_{d-1},j} > 0.$$

故由命题 4.12 知

$$j_1 \in C_{\alpha+1}, j_2 \in C_{\alpha+2}, \cdots, j_{d-1} \in C_{\alpha+d-1}.$$

所以

$$j \in C_{\alpha+d} = C_\alpha.$$

可见 $i \in C_\alpha, j \overline{\in} C_\alpha$ 必导出 $\widetilde{p}_{i,j} = 0$. 此即 $\widetilde{P}_\alpha$ 是转移矩阵.

再证 $\widetilde{P}$ on $C_\alpha = \widetilde{P}_\alpha$ 是不可约的. 事实上, 任取 $i, j \in C_\alpha$, 由 $E$ 的不可约性知: 必存在 $m$, 使 $p_{i,j}^{(m)} > 0$. 又因为 $i$ 和 $j$ 同属于一个循环类 $C_\alpha$, 故必有 $m = rd$, 所以

$$\widetilde{p}_{i,j}^{(r)} = p_{i,j}^{(rd)} = p_{i,j}^{(m)} > 0,$$

此即 $\widetilde{P}_\alpha$ 是不可约的.

最后证明 $\widetilde{P}_\alpha^{\bullet} = \widetilde{P}$ on $C_\alpha$ 是无周期的. 任取 $i \in C_\alpha$, 由命题 4.8 (3) 知: 必存在 $m_0$ 使

$$p_{i,i}^{(m_0 d)} > 0, \quad p_{i,i}^{((m_0+1)d)} > 0.$$

即

$$\widetilde{p}_{i,i}^{(m_0)} > 0, \quad \widetilde{p}_{i,i}^{(m_0+1)} > 0.$$

所以

$$\text{G.C.D.}\{m \geqslant 1 : p_{i,i}^{(m)} > 0\} = 1,$$

此即 $\widetilde{P}_\alpha$ 是无周期. 定理证毕.

由定理 4.2 看出, 任一 $P$ 链的状态空间 $E$ 总可分解成

$$E = N\bigcup R = N\bigcup(R^0\bigcup R^+) = N\bigcup\left(\bigcup_m R_m^0\right)\bigcup\left(\bigcup_n R_n^+\right).$$

由于 $P$ on $R_m^0 \overset{\text{def.}}{=\!=} P_{0,m}$ 是不可约的 ($P$ on $R_n^+ \overset{\text{def.}}{=\!=} P_{+,n}$ 亦然), 所以若令 $R_m^0(R_n^+)$ 之周期为 $d^0(m)(d^+(n))$, 则由定理 4.3, $R_m^0(R_n^+)$ 恰可分解成 $d^0(m)(d^+(n))$ 个循环类:

$$R_m^0 = \bigcup_{t=1}^{d^0(m)} R_m^0(t) \quad \left(R_n^+ = \bigcup_{t=1}^{d^+(n)} R_n^+(t)\right),$$

$P_{0,m}^{d^0(m)}$ 和 $P_{+,n}^{d^+(n)}$ 都是不可约无周期的转移矩阵. 而且

$$p_{i,j}^{(r)} > 0, i \in R_m^0(s), \quad j \in R_m^0(t) \Rightarrow t - s = r (\operatorname{mod} d^0(m)),$$

**定理 4.4** 对任意 $P$ 链的状态空间 $E$, 总有

(1) $i \in R \Leftrightarrow \sum\limits_{n=0}^{\infty} p_{i,i}^{(n)} = \infty \Leftrightarrow f_{i,i}^* = 1 \Leftrightarrow g_{i,i} = 1$;

$$\begin{aligned} i \in R^+ &\Leftrightarrow \sum_{n=0}^{\infty} p_{i,i}^{(n)} = \infty \text{ 且 } F_{i,i}'(1) < \infty \\ &\Leftrightarrow f_{i,i}^* = 1 \text{ 且 } F_{i,i}'(1) < \infty \\ &\Leftrightarrow g_{i,i} = 1 \text{ 且 } F_{i,i}'(1) < \infty, \end{aligned}$$

(2) $i \in R$ 且 $i \leftrightsquigarrow j \Rightarrow f_{i,i}^* = f_{j,j}^* = f_{i,j}^* = f_{j,i}^* = g_{i,i} = g_{j,j} = g_{i,j} = g_{j,i} = 1$.

**证** 由命题 3.4、3.7 和 4.3 立即可得此定理.

**附注 4.1** 在本书中, 矩阵 (特别地, 向量) 的大小的比较、极限等都是逐元意义下的. 例如 $Q^{(n)} = (q_{i,j}^{(n)}, i \in E_1, j \in E_2), Q^{(n)} \geqslant Q^{(m)}$ 的意思是 $q_{i,j}^{(n)} \geqslant q_{i,j}^{(m)}$ (对一切 $i \in E_1, j \in E_2$);

$$\lim_{n\to\infty} Q^{(n)} = \left(\lim_{n\to\infty} q_{i,j}^{(n)}, i \in E_1, j \in E_2\right).$$

零矩阵仍用 0 表示. 矩阵有时分块表示, 例如 $Q = (q_{i,j}, i, j \in E), E = E_1\bigcup E_2$, $E_1\bigcap E_2 = \varnothing$, 有时表

$$Q = \begin{matrix} & \begin{matrix} E_1 & E_2 \end{matrix} \\ \begin{matrix} E_1 \\ E_2 \end{matrix} & \begin{pmatrix} Q_{1,1} & Q_{1,2} \\ Q_{2,1} & Q_{2,2} \end{pmatrix} \end{matrix}$$

其中 $Q_{s,t} = (q_{i,j}, i \in E_s, j \in E_t), 1 \leqslant s, t \leqslant 2$.

## §5 遍历性定理

在这一节中, 恒设 $E$ 是可数集, $P=(p_{i,j},i,j\in E)$ 是 $E$ 上的转移阵, $P^n=P^{(n)}=(p_{i,j}^{(n)},i,j\in E)$ 是 $n$ 步转移阵, $(\Omega^*,\mathscr{F}^*,P^i)(i\in E)$ 是如 §2 中所定义的 $P$ 链的概率空间, 本节研究的主要内容是: $\lim\limits_{n\to\infty}p_{i,j}^{(n)}$ 存在的充要条件是什么? 当此极限存在时, 如何求? 有何性质? 以及状态的区分等等.

**引理 5.1** 设 $\{f_n:n\geqslant 1\},\{p_n:n\geqslant 0\}$ 是两个非负实数序列, $\sum\limits_{n=1}^{\infty}f_n\leqslant 1,p_n\leqslant 1,p_0=1$, 且

$$p_n=\sum_{\nu=1}^{n}f_\nu p_{n-\nu}\quad(n\geqslant 1),\tag{5.1}$$

则 G.C.D.$\{n:f_n>0,n\geqslant 1\}$=G.C.D.$\{n:p_n>0,n\geqslant 1\}$.

**证** 令 $s=\min\{n:f_n>0,n\geqslant 1\}$, 则 $s=\min\{n:p_n>0,n\geqslant 1\}$. 再令 $d'_N$=G.C.D.$\{n:f_n>0,s\leqslant n\leqslant N\},d''_N$ =G.C.D.$\{n:p_n>0,s\leqslant n\leqslant N\}(N\geqslant s)$. 显然 $d'_s=s=d''_s$. 设 $d'_N=d''_N$, 推证: $d'_{N+1}=d''_{N+1}$ 事实上,

$$p_{N+1}=\sum_{\nu=1}^{N}f_\nu p_{N+1-\nu}+f_{N+1},$$

若 $f_{N+1}>0$, 则 $p_{N+1}>0$, 所以由归纳法假设得

$$d'_{N+1}=\text{G.C.D.}\{d'_N:N+1\}=\text{G.C.D.}\{d''_N:N+1\}=d''_{N+1};$$

若 $f_{N+1}=0=p_{N+1}$, 则类似地有 $d'_{N+1}=d''_{N+1}$; 若 $f_{N+1}=0,p_{N+1}>0$, 则存在 $\nu,1\leqslant\nu\leqslant N$, 使 $f_\nu p_{N+1-\nu}>0$. 因此

$$d'_N|\nu,d''_N|(N+1-\nu).$$

从而由 $d'_N=d''_N$ 得 $d'_N|(N+1),d''_N|(N+1)$, 所以 $d'_{N+1}=d'_N=d''_N=d''_{N+1}$. 总之, 恒有 $d'_{N+1}=d''_{N+1}$ (对一切 $N\geqslant s$). 引理证毕.

**定理 5.1** 若 $i$ 是常返状态, 其周期为 $d_i$, 则

$$\lim_{n\to\infty}p_{i,i}^{(nd_i)}=d_i/m_{i,i}\tag{5.2}$$

(其中 $m_{i,i}$ 是平均再现时间, 当 $b=\pm\infty,a$ 是实数时, 此后恒定义 $\dfrac{a}{b}=0$).

**证** 由于 $i$ 是固定的, 简记 $p_{i,i}^{(n)},f_{i,i}^{(n)},d_i,m_{i,i}$ 为 $p_n,f_n,d,m$. 显然 $\{f_n:n\geqslant 1\},\{p_n:n\geqslant 0\}$ 满足引理 5.1 的条件. 令 $r_n=\sum\limits_{\nu=n+1}^{\infty}f_\nu(n\geqslant 0)$, 则

$$r_0p_n=p_n=\sum_{\nu=1}^{n}f_\nu p_{n-\nu}=\sum_{\nu=1}^{n}(r_{\nu-1}-r_\nu)p_{n-\nu}.$$

从而 $\sum\limits_{\nu=0}^{n} r_\nu p_{n-\nu} = \sum\limits_{\nu=0}^{n-1} r_\nu p_{n-1-\nu}$ 不依赖 $n \geqslant 0$. 但 $r_0 p_0 = 1$, 所以

$$\sum_{\nu=0}^{n} r_\nu p_{n-\nu} \equiv 1 \quad (n \geqslant 0). \tag{5.3}$$

令 $\lambda = \lim\limits_{n\to\infty} \sup\, p_{nd} = \lim\limits_{n\to\infty} \sup\, p_n$. 必有 $\{n_k\}$ 使 $\lim\limits_{k\to\infty} p_{n_k d} = \lambda$. 由 (5.1) 及 $\sum\limits_{n=1}^{\infty} f_n = 1$ 可证;

$$\text{“}f_s > 0, \lim_{k\to\infty} p_{m_k} = \lambda \Rightarrow \lim_{k\to\infty} p_{m_k - s} = \lambda\text{”}.$$

所以, 反复利用上述推理可证: “$t = \sum\limits_{j=1}^{l} c_j s_j, c_j, s_j$ 为正整数,

$$f_{sj} > 0, \quad 1 \leqslant j \leqslant l \Rightarrow \lim_{k\to\infty} p_{n_k d - t} = \lambda\text{”}.$$

由引理 5.1 及 $d$ 之定义得知: 存在 $s_j, f_{s_j} > 0, 1 \leqslant j \leqslant l$, 使 $d = \text{G.C.D.}\{s_1, \cdots, s_l\}$. 所以由数论的一条初等定理可知: 存在 $s_0$ 使 “$s \geqslant s_0 \Rightarrow sd = \sum\limits_{j=1}^{l} c_j s_j$”. 因此, $\lim\limits_{k\to\infty} p_{(n_k - s)d} = \lambda (s \geqslant s_0)$. 以 $n = (n_k - s_0)d$ 代入 (5.3) 并注意 $p_\nu = 0$ (当 $\nu \neq 0 (\text{mod } d)$), 有

$$\sum_{\nu=0}^{n_k - s_0} r_{\nu d} p_{(n_k - s_0 - \nu)d} = 1.$$

当 $\sum\limits_{\nu=0}^{\infty} r_{\nu d} < \infty$ 时, 在上式中令 $k \to \infty$ 并应用控制收敛定理 (注意 $\lim\limits_{k\to\infty} p_{(n_k - s)d} = \lambda, s \geqslant s_0$) 可得

$$\lambda \sum_{\nu=0}^{\infty} r_{\nu d} = 1.$$

而当 $\sum\limits_{\nu=0}^{\infty} r_{\nu d} = \infty$ 时易证 $\lambda = 0$. 总之恒有

$$\lambda = \frac{1}{\sum\limits_{\nu=0}^{\infty} r_{\nu d}}.$$

又因为 $f_\nu = 0$ (当 $\nu \neq 0 (\text{mod } d)$), 所以

$$r_{\nu d} = \frac{1}{d} \sum_{j=\nu d}^{\nu d + d - 1} r_j,$$

从而 $\sum\limits_{\nu=0}^{\infty} r_{\nu d} = \dfrac{1}{d} \sum\limits_{j=0}^{\infty} r_j = \dfrac{m}{d}$. 所以 $\lambda = \dfrac{d}{m}$. 若令 $\beta = \lim\limits_{n\to\infty} \inf p_{nd}$, 仿之可证 $\beta = \dfrac{d}{m}$. 定理证毕.

**推论 5.1**　设 $i,j$ 同属于某个周期为 $d$ 的常返类 $R_m^0$ (或 $R_n^+$), $R_m^0 = R_m^0(1)\bigcup\cdots\bigcup R_m^0(d), i\in R_m^0(s), j\in R_m^0(t)$, 则

$$\lim_{n\to\infty} p_{i,j}^{(nd+r)} = \begin{cases} 0, & 若\ r\neq t-s(\mathrm{mod}\ d), \\ \dfrac{d}{m_{j,j}}, & 若\ r=t-s(\mathrm{mod}\ d). \end{cases} \tag{5.4}$$

**证**　当 $r\neq(t-s)(\mathrm{mod}\,d)$ 时, 推论 5.1 的论断显然成立. 当 $r=(t-s)(\mathrm{mod}\,d)$ 时, 由命题 2.3 有

$$\sum_{\nu=1}^{\infty} f_{i,j}^{(\nu d+r)} = 1.$$

再注意: $p_{i,j}^{(\nu)}=0$ (当 $\nu\neq 0(\mathrm{mod}\ d)$), 故

$$p_{i,j}^{(nd+\nu)} = \sum_{\nu=0}^{n} f_{i,j}^{(\nu d+r)} p_{j,j}^{(nd-\nu d)}, \tag{5.5}$$

在 (5.5) 中令 $n\to\infty$ 并应用引理 1.1、定理 5.1 即得 (5.4) 中第 2 式. 推论得证.

对于任何周期为 $d$ 的状态 $j$, 总令

$$f_{i,j}^{*}(r) = \sum_{\substack{n=1\\ n=r(\mathrm{mod}\ d)}}^{\infty} f_{i,j}^{(n)}.$$

**定理 5.2**　(1) 若 $j$ 不是正状态, 则 $\lim\limits_{n\to\infty} p_{i,j}^{(n)} = 0(\forall i\in E)$;
(2) 若 $j$ 是正状态, 周期为 $d_j$, 平均再现时间为 $m_{j,j}$, 且
(a) 若 $i$ 是常返状态但与 $j$ 不在同一类中, 则

$$p_{i,j}^{(n)} = 0 \quad (\forall n\geqslant 1);$$

(b) 若 $i$ 与 $j$ 同属于一个正类, $i\in R_n^+(s), j\in R_n^+(t)$, 则 $p_{i,j}^{(n)}=0$ (当 $n\neq t-s(\mathrm{mod}\ d)$), 而

$$\lim_{n\to\infty} p_{i,j}^{(nd_j+r)} = \frac{d_j}{m_{j,j}} \quad (当\ r=t-s(\mathrm{mod}\ d));$$

(c) 若 $i$ 是暂留状态, 则对一切 $r$, 有

$$\lim_{n\to\infty} p_{i,j}^{(nd_j+r)} = f_{i,j}^{*}(r)\frac{d_j}{m_{j,j}}.$$

**证**　由命题 3.2、4.2 及推论 5.1 即得 (1). 由定理 4.2 得 (2) (a). 由推论 5.1 得 (2) (b). 由 (5.5) 式及引理 3.1 和定理 5.1 得 (2) (c).

**定理 5.3** 对任何 $i,j\in E$, 总有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{n}p_{i,j}^{(\nu)}=\pi_{i,j} \tag{5.6}$$

存在, 其中 $\pi_{i,j}=f_{i,j}^{*}/m_{j,j}$ (当 $j$ 是暂留状态时, 定义 $m_{j,j}=\infty$). 称 $\Pi=(\pi_{i,j},i,j\in E)$ 为 $P$ 的**遍历极限**.

**证** 由定理 5.2 立即可得定理 5.3.

**定理 5.4** $\lim\limits_{n\to\infty}P^{(n)}$ 存在的充要条件是每个正状态的周期都是 1. 如果条件成立, 则 $\lim\limits_{n\to\infty}P^{(n)}=\Pi=(\pi_{i,j},i,j\in E),\pi_{i,j}$ 由定理 5.3 所定义.

**证** 由定理 5.2 和定理 5.3 即得本定理.

下面我们研究 $\Pi$ 的性质及求法.

**命题 5.1** $\forall i\in E,j\in R_n^{+}(R_n^{0}$ 亦类似). 有

$$f_{i,j}^{*}=f_{i,R_n^{+}}^{*}\quad(f_{i,R_n^{+}}^{*}\text{ 之定义见定义 4.4}). \tag{5.7}$$

**证**

$$\begin{aligned}
f_{i,R_n^{+}}^{*}&=\sum_{k=1}^{\infty}P^{i}\left(\begin{bmatrix}1, & \cdots, & k-1, & k\\ (R_n^{+})^{\mathrm c}, & \cdots, & (R_n^{+})^{\mathrm c}, & R_n^{+}\end{bmatrix}\right)\\
&=\sum_{k=1}^{\infty}P^{i}\left(\begin{bmatrix}1, & \cdots, & k-1, & k, & k+1,k+2,\cdots\\ (R_n^{+})^{\mathrm c}, & \cdots, & (R_n^{+})^{\mathrm c}, & R_n^{+}, & j^{\mathrm c},j^{\mathrm c},\cdots\end{bmatrix}\right)\\
&\quad+\sum_{k=1}^{\infty}P^{i}\left(\begin{bmatrix}1, & \cdots, & ,k-1 & ,k\\ (R_n^{+})^{\mathrm c} & ,\cdots, & (R_n^{+})^{\mathrm c}, & R_n^{+}\end{bmatrix},\bigcup_{m=k+1}^{\infty}\begin{bmatrix}m\\ j\end{bmatrix}\right).
\end{aligned} \tag{5.8}$$

用 (2.1) 并注意 $f_{j,j}^{*}=1(j\in R_n^{+})$ 可证 (5.8) 右方第一项为 0, 而第二项等于

$$\begin{aligned}
&P^{i}\left(\bigcup_{k=1}^{\infty}\left(\begin{bmatrix}1, & \cdots, & k-1, & k\\ (R_n^{+})^{\mathrm c}, & \cdots, & (R_n^{+})^{\mathrm c}, & R_n^{+}\end{bmatrix}\bigcap\bigcup_{m=k+1}^{\infty}\begin{bmatrix}m\\ j\end{bmatrix}\right)\right)\\
&\leqslant P^{i}\left(\bigcup_{m=1}^{\infty}\begin{bmatrix}m\\ j\end{bmatrix}\right)=f_{i,j}^{*}.
\end{aligned}$$

所以 $f_{i,R_n^{+}}^{*}\leqslant f_{i,j}^{*}$, 而小于号不能成立. 命题得证.

**命题 5.2** $\Pi=\Pi P=P\Pi=\Pi^{2}$

**证**　用控制收敛定理可得

$$P\Pi = P\left(\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu)}\right) = \lim_{n\to\infty}P\left(\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu)}\right)$$
$$= \lim_{n\to\infty}\left[\left(\frac{1}{n+1}\sum_{\nu=1}^{n+1}P^{(\nu)}\right)\frac{n+1}{n} - \frac{1}{n}P\right] = \Pi.$$

用 Fatou 引理可得

$$\Pi P = \left(\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu)}\right)P \leqslant \lim_{n\to\infty}\left(\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu+1)}\right) = \Pi,$$

令

$$\mathbf{1} = \begin{pmatrix}1\\1\\\vdots\end{pmatrix},$$

则 $\Pi P\mathbf{1} = \Pi\mathbf{1}$, 故上式不等号不可能成立, 所以 $\Pi P = \Pi$.

由 $\Pi = \Pi P$ 得 $\Pi = \Pi P^{(n)} (n \geqslant 0)$, 所以 $\Pi = \Pi\left(\frac{1}{n}\sum\limits_{\nu=1}^{n}P^{(\nu)}\right) (n \geqslant 1)$, 令 $n \to \infty$ 并用控制收敛定理可得

$$\Pi = \lim_{n\to\infty}\Pi\left(\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu)}\right) = \Pi\left(\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{n}P^{(\nu)}\right) = \Pi^2.$$

命题证毕.

**定理 5.5**　记

$$E = N\bigcup R = N\bigcup R^0\bigcup R^+ = N\bigcup\left(\bigcup_m R_m^0\right)\bigcup\left(\bigcup_n R_n^+\right),$$

则 $\Pi$ 具有下述形式:

| | $N$ | $R^0$ | $R_1^+$ | $R_2^+$ | $R_3^+$ | $\cdots$ |
|---|---|---|---|---|---|---|
| $N$ | 0 | 0 | $A(1)$ | $A(2)$ | $A(3)$ | $\cdots$ |
| $R^0$ | 0 | 0 | 0 | 0 | 0 | |
| $R_1^+$ | 0 | 0 | $B(1)$ | 0 | 0 | |
| $R_2^+$ | 0 | 0 | 0 | $B(2)$ | 0 | |
| $R_3^+$ | 0 | 0 | 0 | 0 | $B(3)$ | $\cdots$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |

即 $\pi_{i,j}=0(i\in E,j\in N\bigcup R^0)$, $A(m)=(\pi_{i,j},i\in N,j\in R_m^+)$, $B(n)=\Pi$ on $R_n^+$, $\pi_{i,j}=0(i\overline{\in}(N\bigcup R_n^+),j\in R_n^+)$. 若记 $\pi'(n)=\left(\dfrac{1}{m_{j,j}},j\in R_n^+\right)$ 是以 $\dfrac{1}{m_{j,j}}$ 为分量、定义域为 $R_n^+$ 的行向量, $a(m)=(f^*_{i,R_m^+},i\in N)$ 是以 $f^*_{i,R_m^+}$ 为分量、定义域为 $N$ 的列向量, 则有 $B(n)=\mathbf{1}\pi'(n)$, $A(m)=a(m)\pi'(m)$, 此外还有 $B(n)\mathbf{1}=\mathbf{1}$, 即是 $B(n)$ 是行行一样的每个元素都大于 0 的转移矩阵.

**证** 用定理 5.2 及命题 5.1. 为证此定理, 只需证明 $B(n)\mathbf{1}=\mathbf{1}$ 即可. 事实上, 由命题 5.2, $\Pi=\Pi^2$ 得 $B(n)=B(n)^2$, 即 $\mathbf{1}\pi'(n)=(\mathbf{1}\pi'(n)\mathbf{1}\pi'(n))=\mathbf{1}(\pi'(n)\mathbf{1})\pi'(n)=(\pi'(n)\mathbf{1})(\mathbf{1}\pi'(n))$, 由 $\mathbf{1}\pi'(n)$ 每一元素均为正知 $\pi'(n)\mathbf{1}=1$, 故 $B(n)\mathbf{1}=(\mathbf{1}\pi'(n))\mathbf{1}=\mathbf{1}(\pi'(n)\mathbf{1})=\mathbf{1}$. 定理证毕.

**定义 5.1** 称 $Q=(q_{i,j},i,j\in E)$ 是**准转移矩阵**, 如果 $q_{i,j}\geqslant 0$, $\sum\limits_{k\in E}q_{i,k}\leqslant 1(i,j,\in E)$. 称准转移矩阵 $Q$ 具有 $\Pi$ 结构, 如果 $E$ 可分成不交子集的并: $E=G\bigcup F_1\bigcup F_2\bigcup\cdots$ ($G$ 或 $F_n$ 可为空集), 使 $Q$ 具有下述形状:

| | $G$ | $F_1$ | $F_2$ | $\cdots$ |
|---|---|---|---|---|
| $G$ | 0 | $C(1)q'(1)$ | $C(2)q'(2)$ | $\cdots$ |
| $F_1$ | 0 | $\mathbf{1}q'(1)$ | 0 | |
| $F_2$ | 0 | 0 | $\mathbf{1}q'(2)$ | |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | |

其中 $C(m)$ 是以 $G$ 为定义域、分量为非负实数的列向量, $q'(n)$ 是以 $F_n$ 为定义域、分量为非负实数的行向量, 且 $q'(n)\mathbf{1}=\mathbf{1}$.

显然, 若 $Q$ 具有 $\Pi$ 结构, 则必有 $Q^2=Q$ (注意 $q'(n)\mathbf{1}=1$). 又由定理 5.5, $\Pi$ 是具有 $\Pi$ 结构的 (取 $G=N\bigcup R^0, F_n=R_n^+, q'(n)=\pi'(n), C(m)=\binom{a(m)}{0}$).

**命题 5.3** 若 $Q=(q_{i,j},i,j\in E)$ 是准转移矩阵, 则 $\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{\nu=1}^{n}Q^\nu=\widetilde{\Pi}$ 存在且具有 $\Pi$ 结构 ($Q^\nu$ 是 $Q$ 的 $\nu$ 次幂).

**证** 令 $\Delta$ 是 $E$ 外之一点, $E_\Delta=E\bigcup\{\Delta\}$ ($\{\Delta\}$ 表示由 $\Delta$ 构成的单点集), $b=\mathbf{1}-Q\mathbf{1}$, 则 $Q_\Delta=\begin{pmatrix}Q & b\\ 0 & 1\end{pmatrix}$ 是定义在 $E_\Delta$ 上的转移矩阵. 由定理 5.5 知

$$\lim_{n\to\infty}\frac{1}{n}\sum_{\nu=1}^{n}Q_\Delta^\nu=\lim_{n\to\infty}\begin{pmatrix}\dfrac{1}{n}\sum\limits_{\nu=1}^{n}Q^\nu & b_n\\ 0 & 1\end{pmatrix}$$

存在且有 $\Pi$ 结构, 故 $\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{\nu=1}^{n}Q^\nu$ 存在也有 $\Pi$ 结构.

**命题 5.4** $Q$ 是准转移矩阵且 $Q=Q^2$ 的充要条件是 $Q$ 具有 $\Pi$ 结构.

**证** 充分性由 $\Pi$ 结构之定义即得. 必要性由 $Q=\frac{1}{n}\sum\limits_{\nu=1}^{n}Q^{\nu}$ 及命题 5.3 即得.

**命题 5.5** $f^*_{i,J}(i\in E, J\subset E)$ 具有下列性质:

(1) $f^*_{i,\phi}=0, f^*_{i,E}=1(i\in E)$;

(2) $J_1\subset J_2\Rightarrow f^*_{i,J_1}\leqslant f^*_{i,J_2}$;

(3) $f^*_{i,J}$ 对 $J$ 具有完全半可加性;

(4) $\{J_\nu\}$ 是不交封闭集 $\Rightarrow f^*_{i,\bigcup\limits_\nu J_\nu}=\sum\limits_\nu f^*_{i,J_\nu}$;

(5) 若 $J\subset E$ 固定, 则或 $f^*_{i,J}=1$ (对一切 $i\in E$), 或 $\inf\limits_{i\in J}f^*_{i,J}=0$.

**证** (1)、(2) 显然成立.

(3)
$$f^*_{i\bigcup\limits_\nu J_\nu}=P^i\left(\bigcup_n\left[\begin{matrix}n\\ \bigcup\limits_\nu J_\nu\end{matrix}\right]\right)=P^i\left(\bigcup_\nu\bigcup_n\left[\begin{matrix}n\\ J_\nu\end{matrix}\right]\right), \tag{5.9}$$

所以 $f^*_{i,\bigcup\limits_\nu J_\nu}\leqslant\sum_\nu f^*_{i,J_\nu}$.

(4) 若 $\{J_\nu\}$ 不交封闭, 则当 $\mu\neq\nu$ 时,

$$P^i\left(\left(\bigcup_n\left[\begin{matrix}n\\ J_\nu\end{matrix}\right]\right)\cap\left(\bigcup_n\left[\begin{matrix}n\\ J_\mu\end{matrix}\right]\right)\right)=0,$$

故由 (5.9) 即得 (4).

(5) 由 (2.1) 得

$$\begin{aligned}
1-f^*_{i,J}&=P^i\left(\bigcap_{n=1}^{\infty}\left[\begin{matrix}n\\ J^c\end{matrix}\right]\right)\\
&=\sum_{j\in J^c}P^i\left(\left[\begin{matrix}1, & \cdots, & n-1, & n\\ J^c, & \cdots, & J^c, & j\end{matrix}\right]\right)P^j\left(\left[\begin{matrix}1, & 2, & \cdots\\ J^c, & J^c, & \cdots\end{matrix}\right]\right)\\
&=\sum_{j\in J^c}P^i\left(\left[\begin{matrix}1, & \cdots, & n-1, & n\\ J^c, & \cdots, & J^c, & j\end{matrix}\right]\right)(1-f^*_{j,J})\\
&\leqslant P^i\left(\bigcap_{k=1}^{n}\left[\begin{matrix}k\\ J^c\end{matrix}\right]\right)\left(1-\inf_{j\in J^c}f^*_{j,J}\right),
\end{aligned}$$

令 $n\to\infty$ 即得: $1-f^*_{i,J}\leqslant(1-f^*_{i,J})(1-\inf\limits_{i\in J}f^*_{j,J})$, (5) 得证.

令 $\gamma = \Pi \mathbf{1}$, 则

$$\gamma = \begin{matrix} N \\ R^0 \\ R^+ \end{matrix} \begin{pmatrix} \sum_m a(m) \\ 0 \\ \mathbf{1} \end{pmatrix}, \tag{5.10}$$

即是 $\gamma$ on $N = \sum_m a(m), \gamma$ on $R^0 = 0, \gamma$ on $R^+ = \mathbf{1}, a(m), \Pi$ 之定义见定理 5.5.

本书中恒用 $(m)_J$ 表示定义域为 $J$、分量有界的列向量全体 $(l)_J$ 表示定义域为 $J$、诸分量之绝对值之和收敛的行向量的全体. 若 $J = E$, 则简记 $(m)_E$、$(l)_E$ 为 $(m)$、$(l)$. 用 $e_i(e_i')$ 表示对应于 $i$ 的分量为 1、其他分量为 0 的列 (行) 向量, 其维数视需要而定.

**命题 5.6** $\{e_i'\gamma : i \in E\}$ 或无 0 或有无穷多个 0.

**证** 若 $R^0$ 不空, 则 $P$ on $R^0$ 是转移矩阵, 而 $\lim_{n\to\infty}(P^{(n)} \text{ on } R^0) = 0$, 故 $R^0$ 必为无穷集, 由 (5.10) 知 $\{e_i'\gamma : i \in E\}$ 有无穷多个 0.

若 $R^0$ 是空集, 再设 $\{i : e_i'\gamma = 0, i \in E\} \neq \varnothing$. 由 (5.10) 和命题 5.5(4) 及定理 5.5 知: 当 $i \in N$ 时有 $e_i'\gamma = \sum_m f_{i,R_m^+}^* = f_{i,R^+}^*$. 所以由 (5.10) 及 $R^0$ 是空集知:

$$\begin{aligned} \varnothing \neq \{i : e_i'\gamma = 0, i \in E\} &= \{i : e_i'\gamma = 0, i \overline{\in} R^+\} \\ &= \{i : f_{i,R^+}^* = 0, i \overline{\in} R^+\} = \widehat{R}^+(R^+)^c \end{aligned}$$

($\widehat{R}^+$ 之定义见定义 4.4). 若 $R^+$ 是空集, 则 $e_i'\gamma \equiv 0 (i \in E)$, 命题 5.6 成立, 若 $R^+$ 非空, 用命题 4.4 知 $\widehat{R}^+(R^+)^c$ 为封闭集, 故 $P$ on $(\widehat{R}^+(R^+)^c)$ 是转移矩阵, 若注意 $\widehat{R} + (R^+)^c \subset N$ (因 $R^0$ 是空集) 知

$$\lim_{n\to\infty}(P \text{ on } \widehat{R}^+(R^+)^c)^n = \lim_{n\to\infty}(P^{(n)} \text{ on } \widehat{R}^+(R^+)^c) = 0,$$

所以 $\widehat{R}^+(R^+)^c$ 是无穷集, 命题 5.6 证毕.

**命题 5.7** 下列陈述等价:

(1) $\Pi$ 是转移矩阵;

(2) $R^0 = \varnothing, f_{i,R^+}^* = 1 (i \in N)$;

(3) $\inf\{e_i'\gamma : i \in E\} > 0$.

**证** 由 (5.10) 及 $e_i'\gamma = f_{i,R^+}^* (i \in N)$ 知 (1) $\Leftrightarrow$ (2). 而 (1) $\Rightarrow$ (3) 显然, 下面证明 (3) $\Rightarrow$ (2). 若 (3) 成立, 则 $R^0 = \varnothing$,

$$\inf_{i \in N} f_{i,R^+}^* = \inf_{i \in (R^+)^c} f_{i,R^+}^* > 0,$$

用命题 5.5 (5) 得 $f_{i,R^+}^* = 1$ (对一切 $i \in E$), 故 (2) 成立.

由命题 5.1、5.7 及定理 5.3 知: "$\Pi = \mathbf{1}\pi', \pi'\mathbf{1} = 1, \pi' \geqslant 0 \Leftrightarrow R^0 = \varnothing$, 恰有一个正类, 且 $f^*_{i,R_1^+} = 1(i \in N)$."

**定义 5.2** 称转移矩阵 $P$ (或对应的 $P$ 链) 是**无耗损的**, 如果 $\Pi$ 是转移矩阵; 反之称 $P$ 是**耗损的**.

**引理 5.2** 设 $\alpha'(n) = (\alpha_0(n), \alpha_1(n), \cdots)$, $\alpha' = (\alpha_0, \alpha_1, \cdots)$, $\alpha'(n) \geqslant 0$, $\alpha'(n)\mathbf{1} = 1$ $(n \geqslant 1)$,

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \end{pmatrix},$$

$\lim\limits_{n\to\infty} \alpha'(n) = \alpha'$. 若下列两条件之一成立:

(1) $\lim\limits_{k\to\infty} \sup\limits_{n\geqslant 1} \sum\limits_{i=k}^{\infty} \alpha_i(n) = 0$; (2) $\alpha'(n)\beta \leqslant c(n \geqslant 1), \lim\limits_{k\to\infty} \beta_k = \infty, \beta \geqslant 0$, 则 $\alpha'\mathbf{1} = 1$.

**证** 设 (1) 成立, 则对任给 $\varepsilon > 0$, 存在 $k_0$, 使

$$\sup_{n\geqslant 1} \sum_{i=k_0}^{\infty} \alpha_i(n) < \frac{\varepsilon}{2}.$$

再用 $\lim\limits_{n\to\infty} \alpha'(n) = \alpha'$ 得: 存在 $N_0$ 使 $|\alpha_i(N_0) - \alpha_i| < \dfrac{\varepsilon}{2k_0}(i = 1, 2, \cdots, k_0)$. 所以

$$\begin{aligned}\alpha'\mathbf{1} &\geqslant \sum_{i=1}^{k_0} \alpha_i \geqslant \sum_{i=1}^{k_0} \left(\alpha_i(N_0) - \frac{\varepsilon}{2k_0}\right) \\ &\geqslant 1 - \sum_{i=k_0}^{\infty} \alpha_i(N_0) - \frac{\varepsilon}{2} > 1 - \varepsilon.\end{aligned}$$

由 $\varepsilon$ 之任意性得 $\alpha'\mathbf{1} \geqslant 1$, 再用 Fatou 引理得 $\alpha'\mathbf{1} = 1$.

设 (2) 成立. 则 $c \geqslant \sum\limits_{i=k}^{\infty} \alpha_i(n)\beta_i \geqslant (\inf\limits_{i\geqslant k} \beta_i) \sum\limits_{i=k}^{\infty} \alpha_i(n), \sup\limits_{n\geqslant 1} \sum\limits_{i=k}^{\infty} \alpha_i(n) \leqslant c / \inf\limits_{i\geqslant k} \beta_i$, 由 (2) 即得 (1). 引理 5.2 证毕.

**定理 5.6** 设 $E = \{0, 1, 2, \cdots\}$,

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \end{pmatrix} \geqslant 0,$$

$\lim\limits_{k\to\infty} \beta_k = \infty, P\beta \leqslant \beta$, 则 $P$ 是无耗损的.

**证** 任取 $i \in E$ 固定, 令 $\alpha'(n) = e_i'\left(\frac{1}{n}\sum_{\nu=1}^{n} P^{\nu}\right), \alpha' = e_i'\Pi$ 用引理 5.2 (2) 即得 $\alpha'\mathbf{1} = 1$, 故 $\Pi\mathbf{1} = 1$. 定理得证.

**定理 5.7** 设 $E = \{0, 1, 2, \cdots\}$,

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \end{pmatrix} \geqslant 0,$$

$\alpha$ 是正实数, $P\beta < \infty, e_i'P\beta \leqslant e_i'\beta - \alpha$ (除去 $E$ 中有限个 $i$), 则 $P$ 是无耗损的.

**证** 由假设知: 存在 $b > 0$ 使 $P\beta \leqslant \beta + b\mathbf{1}$, 故 $P^{(n)}\beta < \infty (n \geqslant 1)$. 再用假设, 存在 $c > 0$ 及一个列向量 $\alpha, \alpha$ 只有有限个分量为 1 其他为 0, 使 $P\beta \leqslant \beta - \alpha\mathbf{1} + c\alpha$. 由 $P\beta + \alpha\mathbf{1} \leqslant \beta + c\alpha$ 得

$$\left(\frac{1}{n}\sum_{\nu=1}^{n} P^{(\nu)}\right) P\beta + \alpha\mathbf{1} \leqslant \left(\frac{1}{n}\sum_{\nu=1}^{n} P^{(\nu)}\right)\beta + c\left(\frac{1}{n}\sum_{\nu=1}^{n} P^{(\nu)}\right)\alpha.$$

注意 $P^{(n)}\beta < \infty (n \geqslant 1)$ 可得:

$$\alpha\mathbf{1} \leqslant \frac{1}{n}P\beta - \frac{1}{n}P^{(n+1)}\beta + c\left(\frac{1}{n}\sum_{\nu=1}^{n} P^{(\nu)}\right)\alpha.$$

令 $n \to \infty$ 即得 $\alpha\mathbf{1} \leqslant c\Pi\alpha \leqslant c\Pi\mathbf{1} = c\gamma$. 故 $\inf\{e_i'\gamma : i \in E\} \geqslant \frac{a}{c} > 0$, 所以由命题 5.7 知 $P$ 是无耗损的.

**定义 5.3** 称分布行 $\beta'(\beta' \geqslant 0, \beta'\mathbf{1} = 1)$ 是转移矩阵 $P$ (或对应的 $P$ 链) 的**平稳分布**或**不变测度**或**谱测**, 如果 $\beta'P = \beta'$.

称 "$x'P = x', x' \in (l), x' \geqslant 0$" 为**左方程**.

**定理 5.8** 下列 4 条陈述等价:

(1) 方程 "$x'P = x', x' \in (l)$" 有非 0 解;

(2) 有正状态存在;

(3) 左方程有非 0 解;

(4) $P$ 有平稳分布存在.

**证** 先考察 "$x'P = x', x' \in (l)$" 的通解. 用控制收敛定理可得: 右 $x'P = x', x' \in (l)$, 则 $x'\Pi = x', x' \in (l)$. 再用定理 5.5 得: $x'$ on $(N \bigcup R^0)$ 为 0, 若记

$x'$ on $R_n^+$ 为 $s_n'$, 则再用定理 5.5 有

$$
\begin{aligned}
x' &= (0,0,s_1',s_2',\cdots) = (0,0,s_1',s_2',\cdots)\Pi \\
&= (0,0,s_1',s_2',\cdots)\begin{pmatrix} 0 & 0 & a(1)\pi'(\mathbf{1}) & a(2)\pi'(2) & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \mathbf{1}\pi'(1) & 0 & \cdots \\ 0 & 0 & 0 & \mathbf{1}\pi'(2) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \\
&= (0,0,s_1\pi'(1),s_2\pi'(2),\cdots),
\end{aligned}
$$

即 $x'$ on $(N\bigcup R^0)=0, x'$ on $R_n^+ = s_n\pi'(n), s_n = s_n'\mathbf{1}, \sum\limits_n |s_n| < \infty$.

反之, 任给满足上述要求的 $x'$, 必满足 $x'P = x', x' \in (l)$, 故 "$x'P = x', x' \in (l)$" 的通解为

$$
\begin{array}{l}
\quad\;\; N \quad R^0 \quad\;\; R_1^+ \qquad\quad R_2^+ \\
\begin{cases} x' = (0,\ 0,\ s_1\pi'(1),\ s_2\pi'(2),\cdots), \\ \sum\limits_n |s_n| < \infty. \end{cases}
\end{array}
\tag{5.11}
$$

(1) $\Rightarrow$ (2). 由通解的形状即得.

(2) $\Rightarrow$ (3). 在 (5.11) 中取 $s_1 > 0, s_i = 0 (i > 1)$ 即为左方程之非 0 解.

(3) $\Rightarrow$ (4). 设 $x'$ 是左方程之非 0 解, 取 $\beta' = \dfrac{1}{x'\mathbf{1}}x'$ 即为 $P$ 的平稳分布.

(4) $\Rightarrow$ (1), 显然.

**定理 5.9** *若 $P$ 是不可约的, 则*

*(1) 左方程只有零解 $\Rightarrow \Pi = 0$;*

*(2) 左方程有非零解 $\Rightarrow \Pi = \mathbf{1}\pi'$, 其中 $\pi' = \dfrac{1}{x'\mathbf{1}}x', x'$ 是左方程的唯一非零解 (常因子除外), $E = R_1^+$, 且 $\pi'$ 是 $P$ 的唯一的平稳分布.*

**证** (1) 由左方程只有零解及定理 5.8 知: 无正状态, 故 $\Pi = 0$.

(2) 由左方程有非零解知有正状态存在, 而 $P$ 不可约, 故 $E = R_1^+$. 故左方程之非 0 解为 $x' = s_1\pi'(1), s_1 > 0$, 而 $\Pi = \mathbf{1}\pi'(1)$, 故

$$
\Pi = \mathbf{1}\left(\frac{1}{x'\mathbf{1}}x'\right).
$$

由 $\Pi P = \Pi$ 知: $\pi' = \pi'(1)$ 是 $P$ 的唯一的平稳分布.

对一般的转移阵 $P$ 而言 (可能是可约的), 解左方程, 其通解必为 0 的地方对应的状态就是 $N\bigcup R^0$, 其余的地方就是 $R_1^+, R_2^+, \cdots$, 而且 $\pi'(n)$ 也得出了. 根

据定理 5.5, $\Pi$ 的结构, 只需求出 $a(m)$, $\Pi$ 就完全求出了. 下述定理就回答了 $a(m)$ 的求法.

**定理 5.10** 对任何转移矩阵 $P$ 而言,

$$\begin{matrix} N \\ R^0 \\ \bigcup\limits_{s<m} R_s^+ \\ R_m^+ \\ \bigcup\limits_{s>m} R_s^+ \end{matrix} \begin{pmatrix} a(m) \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix}$$

是

$$《P \begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix} = \begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix}, 0 \leqslant y \leqslant 1》$$

的最小解.

**证** 由 $P\Pi = \Pi$ 并应用定理 5.5 $\Pi$ 的结构再注意 $\pi'(m) > 0$ 即可发现它是解. 再证最小性. 任取上述方程的一个解

$$\begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix}, \quad 0 \leqslant y \leqslant \mathbf{1},$$

必有

$$\begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix} = \left( \frac{1}{n} \sum_{\nu=1}^{n} P^{\nu} \right) \begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix},$$

令 $n \to \infty$ 并用 Fatou 引理及定理 5.5 $\Pi$ 的结构可得

$$\begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix} \geqslant \Pi \begin{pmatrix} y \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix} \geqslant \Pi \begin{pmatrix} 0 \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix} = \begin{pmatrix} a(m) \\ 0 \\ 0 \\ \mathbf{1} \\ 0 \end{pmatrix}.$$

至此, $\Pi$ 的求法及 $N \bigcup R^0$ 与 $\bigcup\limits_n R_n^+$ 的区分已经解决. 本节最后一个问题就是要区分 $N$ 与 $R^0$.

**命题 5.8** 任取 $J \subset E, J \neq \varnothing, A = P$ on $J$, 则

(1) $\lim\limits_{n\to\infty} A^n\mathbf{1} = a$ 存在;

(2) $Aa = a$, 且 $a$ 是《$Ay = y, 0 \leqslant y \leqslant \mathbf{1}$》的最大解;

(3) $a_i = e_i'a = 1 - f_{i,J}^*(i \in J)$;

(4) 令 $\widetilde{a}_i = a_i(i \in J), \widetilde{a}_i = 0(i \in J^{\mathrm{c}}), \widetilde{a}(\widetilde{a}_i, i \in E)$, 则 $e_i'(P\widetilde{a}) = 1 - f_{i,J^{\mathrm{c}}}^*(i \in E)$.

**证** 任取 $i \in J$, 则由测度 $P^i$ 的定义有

$$e_i'(A^n\mathbf{1}) = P^i\left(\bigcap_{k=1}^n \begin{bmatrix} k \\ J \end{bmatrix}\right),$$

故 (1)、(3) 得证. 再证 (2). 用控制收敛定理得 $Aa = A\lim\limits_{n\to\infty} A^n\mathbf{1} = \lim\limits_{n\to\infty} A^{n+1}\mathbf{1} = a$. 若 $Ay = y, 0 \leqslant y \leqslant \mathbf{1}$, 则 $y = A^n y \leqslant A^n\mathbf{1}$, 从而 $y \leqslant \lim\limits_{n\to\infty} A^n\mathbf{1} = a$, (2) 得证. 最后证 (4). 用 (3) 有

$$\begin{aligned} 1 - f_{i,\overline{J}}^* &= P^i\left(\bigcap_{n=1}^\infty \begin{bmatrix} n \\ J \end{bmatrix}\right) = \sum_{j\in J} p_{i,j} P^j \left(\bigcap_{n=1}^\infty \begin{bmatrix} n \\ J \end{bmatrix}\right) \\ &= \sum_{j\in J} p_{i,j}(1 - f_{i,J^{\mathrm{c}}}^*) = \sum_{j \in J} p_{i,j}(e_j'a) \\ &= \sum_{j\in E} p_{i,j}\widetilde{a}_j = e_i'(P\widetilde{a}). \end{aligned}$$

**命题 5.9** 下列陈述等价:

(1) $f_{i,J^{\mathrm{c}}}^* \equiv 1(i \in E)$;

(2) $f_{i,J^{\mathrm{c}}}^* = 1(i \in J)$;

(3) $a = 0$ ($a$ 之定义见命题 5.8);

(4)《$Ay = y, y \geqslant 0, y \in (m)$》只有 0 解.

**证** (1) $\Rightarrow$ (2) 显然. (2) $\Rightarrow$ (3), 只需注意 $e_i'a = \mathbf{1} - f_{i,J^{\mathrm{c}}}^*(i \in J)$. (3) $\Rightarrow$ (4), 设 $y$ 为方程之解, 不妨令 $Ay = y, 0 \leqslant y \leqslant \mathbf{1}$, 由命题 5.8 (2) 有 $0 \leqslant y \leqslant a$, 而今

设 $a=0$, 故 $y=0$, (4) 成立. (4) $\Rightarrow$ (1), 由 (4) 成立, 再用命题 5.8 (2) 得 $a=0$, 从而 $P\tilde{a}=0$. 但 $e_i'P\tilde{a}=1-f_{i,\overline{J}}^*(i\in E)$, 所以 $f_{i,J^c}^*\equiv 1(i\in E)$.

**推论 5.2** 设 $A$ 为状态 $j$ 之余阵, 即是 $A=P$ on $(E-\{j\})$, 则下列陈述等价:

(1) $f_{i,j}^*\equiv 1\ (i\in E)$;

(2) $f_{i,j}^*=1\ (i\neq j)$;

(3) $a=0\ (a=\lim\limits_{n\to\infty}A^n\mathbf{1})$;

(4)《$Ay=y,y\geqslant 0,y\in(m)$》只有 0 解.

**定理 5.11** 任取状态 $j$, 设 $A$ 为 $j$ 之余阵, 则

(1)《$Ay=y,y\geqslant 0,y\in(m)$》只有 0 解 $\Rightarrow j$ 是常返状态;

(2)《$Ay=y,y\geqslant 0,y\in(m)$》有非 0 解, 且 $P$ 不可约, 则 $j$ 是暂留状态.

**证** 由命题 5.9 推论 5.2 中 (1) $\Leftrightarrow$ (4) 即得本定理的 (1). 再证 (2). 若 $P$ 不可约, $j$ 是常返状态, 则 $f_{i,j}^*\equiv 1(i\in E)$, 再用推论 5.2 知《$Ay=y,y\geqslant 0,y\in(m)$》只有 0 解,(2) 证毕.

我们称《$Py=y,y\geqslant 0,y\in(m)$》为右方程. 如 $A$ 为 $j$ 之余阵, 即 $A=P$ on $(E-\{j\})$, 则称《$Ay=y,y\geqslant 0,y\in(m)$》为 $j$ 右方程, 引进

条件 (LC); 左方程有非 0 解;

条件 $(\mathrm{RC})_j$ : $j$ 右方程有非 0 解;

条件 (RC): 右方程有非 0 解;

$\overline{(\mathrm{LC})}((\overline{\mathrm{RC}}),\overline{(\mathrm{RC})_j})$ 分别表示 (LC)((RC),$(\mathrm{RC})_j$) 之逆条件.

**定理 5.12** 设 $P$ 不可约, 则

(1) (LC) $\Leftrightarrow E=R_1^+\Rightarrow(\overline{\mathrm{RC}})_j$ (对一切 $j\in E$);

(2) $\overline{(\mathrm{LC})},\overline{(\mathrm{RC})}_j$ (对某个 $j\in E$) $\Leftrightarrow\overline{(\mathrm{LC})},\overline{(\mathrm{RC})}_j$ (对一 $j\in E$) $\Leftrightarrow E=R_1^0$;

(3) $(\mathrm{RC})_j$ (对某个 $j\in E$) $\Leftrightarrow(\mathrm{RC})_j$ (对一切 $j\in E)E=N$.

**证** 由定理 5.11 及左方程之通解的形式即得定理 5.12.

**定理 5.13** 下列陈述等价:

(1) $E$ 不是一个常返类;

(2)《$Py\leqslant y,y\geqslant 0,y\in(m)$》有非常向量解;

(3)《$Py\leqslant y,y\geqslant 0$》有非常向量解.

(常向量者, 即诸分量相等之向量.)

**证** (1) $\Rightarrow$ (2). 先设 $E$ 有暂留状态. 不妨令 $E=\{0,1,\cdots\}$,0 是暂留状态.

取

$$\widetilde{y}=\begin{pmatrix}1,\\ f_{1,0}^{*}\\ f_{2,0}^{*}\\ \vdots\end{pmatrix},$$

由命题 1.5 (3) 有

$$P\widetilde{y}=\begin{pmatrix}f_{0,0}^{*}\\ f_{1,0}^{*}\\ \vdots\end{pmatrix}\leqslant\widetilde{y}.$$

显然 $f_{i,0}^{*}(i\geqslant 1)$ 不能全是 $\mathbf{1}$ (否则由命题 5.9 的推论 5.2 得 $f_{0,0}^{*}=1$, 与 0 是暂留状态矛盾), 而且由 $f_{0,0}^{*}<1$ 知 $P\widetilde{y}\neq\widetilde{y}$. 总之, 当 $E$ 有暂留状态时,

$$《\,Py\leqslant y, y\geqslant 0, y\in(m), Py\neq y\,》$$

有非常向量解 $\widetilde{y}$.

再设 $E$ 无暂留状态, 由 (1) 成立和 $E$ 至少含两个常返类, 不妨令 $R_1^{+}, R_2^{+}$ 非空, 取向量 $y^{*}$ 满足: $e_i'y^{*}=c_1\geqslant 0(i\in R_1^{+}), e_i'y^{*}=c_2\geqslant 0(j\in R_2^{+}), c_1\neq c_2, e_k'y^{*}=0(k\in E-R_1^{+}\bigcup R_2^{+})$, 则 $Py^{*}=y^{*}$. 故

$$《\,Py\leqslant y, y\geqslant 0, y\in(m)\,》$$

有非常向量解 $y^{*}$, (1) $\Rightarrow$ (2) 证毕.

(2) $\Rightarrow$ (3) 显然.

(3) $\Rightarrow$ (1). 令 $E=\{0,1,2,\cdots\}, \widetilde{P}=(\widetilde{p}_{i,j}, i,j\in E), p_{0,0}=1, \widetilde{p}_{0,j}=0(j\geqslant 1), \widetilde{p}_{i,j}=p_{i,j}(i\geqslant 1, j\in E), \widetilde{\Pi}=\lim\limits_{n\to\infty}\dfrac{1}{n}\sum\limits_{\nu=1}^{n}\widetilde{P}^{(\nu)}, A=P$ on $(E-\{0\})=\widetilde{P}$ on $(E-\{0\})$, 则

$$\widetilde{P}^{(n)}e_0=\widetilde{P}^{(n)}\left(\begin{pmatrix}1\\ \mathbf{1}\end{pmatrix}-\begin{pmatrix}0\\ \mathbf{1}\end{pmatrix}\right)=\begin{pmatrix}1\\ \mathbf{1}\end{pmatrix}-\begin{pmatrix}0\\ A^{n}\mathbf{1}\end{pmatrix}.$$

由命题 5.8(1) 及 (3) 得 (取 $\{0\}$ 为那儿的 $J^{c}$):

$$\lim_{n\to\infty}\widetilde{P}^{(n)}e_0=\widetilde{y}\ \ (\widetilde{y}\text{ 之定义见前})$$

从而 $\widetilde{\Pi}e_0=\widetilde{y}$. 下面证明: $\widetilde{y}$ 是《$Py\leqslant y, y\geqslant 0, e_0'y=\mathbf{1}$》的最小解. $\widetilde{y}$ 是解在 (1) $\Rightarrow$ (2) 的证明中已证. 再证最小性, 若 $y\geqslant 0, e_0'y=1, y\geqslant Py$, 则 $y\geqslant\widetilde{P}y\geqslant\cdots\geqslant\widetilde{P}^{(n)}y$, 从而

$$y\geqslant\left(\frac{1}{n}\sum_{\nu=1}^{n}\widetilde{P}^{(\nu)}\right)y,$$

令 $n \to \infty$ 用 Fatou 引理得 $y \geqslant \widetilde{\Pi} y \geqslant \widetilde{\Pi} e_0 = \widetilde{y}$. 最小性证毕. 现在设 (3) 成立, 即存在 $y \geqslant 0, y \geqslant Py, y$ 是非常向量, 不妨令 $e_0'y > e_n'y$, 又不妨令 $e_0'y = 1 > e_1'y$ (否则以 $e_0'y$ 除此向量). 由 $\widetilde{y}$ 之最小性知 $\widetilde{y} \leqslant y$, 更有 $f_{1,0}^* = e_1'\widetilde{y} \leqslant e_1'y < 1$, 所以 $E$ 不是一个常返类. 定理证毕.

**定理 5.14** 下列陈述等价;

(1) $E$ 有暂留状态;

(2) 《$Py \leqslant y, y \geqslant 0, y \in (m), Py \neq y$》有解;

(3) 《$Py \leqslant y, y \geqslant 0, Py \neq y$》有解.

**证** (1) $\Rightarrow$ (2) 在定理 5.13 的 (1) $\Rightarrow$ (2) 中已证.

(2) $\Rightarrow$ (3) 显然成立.

(3) $\Rightarrow$ (1). 设 $y \geqslant Py, y \geqslant 0, Py \neq y$. 作

$P^* = \text{diag}([1+e_i'y]^{-1}, i \in E)P\text{diag}([1+e_i'y], i \in E)$, 其中 $\text{diag}(q_i, i \in E)$ 表示对角矩阵, 主对角线上对应于 $i$ 的元为 $q_i$) 则 $e_i'P^{(\nu)}e_i = e_i'(P^*)^\nu e_i (\nu \geqslant 0, i \in E)$, 且

$$\begin{aligned} P^*\mathbf{1} &= \text{diag}([1+e_i'y]^{-1}, i \in E)P(\mathbf{1}+y) \\ &\leqslant \text{diag}([1+e_i'y]^{-1}, i \in E)(\mathbf{1}+y) \\ &= \mathbf{1}, \end{aligned}$$

由 $Py \leqslant y, Py \neq y$ 知上式中等号不能成立, 故

$$\mathbf{1} - P^*\mathbf{1} \geqslant 0, \quad \mathbf{1} - P^*\mathbf{1} \neq 0,$$

所以存在 $i_0 \in E$ 及正数 $c$ 使 $\mathbf{1} - P^*\mathbf{1} \geqslant ce_{i_0}$. 因此

$$\mathbf{1} \geqslant \sum_{\nu=0}^{n=1}(P^*)^\nu(\mathbf{1} - P^*\mathbf{1}) \geqslant c\sum_{\nu=0}^{n-1}(P^*)^\nu e_{i_0}.$$

特别地, 有

$$1 \geqslant e_{i_0}'c\sum_{\nu=0}^{n-1}(P^*)^\nu e_{i_0} = c\sum_{\nu=0}^{n=1} e_{i_0}'P^{(\nu)}e_{i_0} = c\sum_{\nu=0}^{n=1} p_{i_0,i_0}^{(\nu)} \quad (n \geqslant 1)$$

所以 $\sum\limits_{\nu=0}^{\infty} p_{i_0,i_0}^{(\nu)} < \infty$, 即 $i_0$ 是暂留状态.

**定理 5.15** 设 $\widetilde{P}$ 如定理 5.13 所定义, 若 $\widetilde{P}$ 是无耗损的, 则有常返状态.

**证** $E = \{0, 1, 2, \cdots\}$, 若 $1, 2, \cdots$ 都是 $P$ 的暂留状态, 则

$$\sum_{n=0}^{\infty} p_{i,i}^{(n)} < \infty \quad (i \geqslant 1).$$

又由 $\widetilde{P}$ 之定义知 $\widetilde{p}_{i,i}^{(n)} \leqslant p_{i,i}^{(n)}(i \geqslant 1, n \geqslant 0)$, 所以 $1, 2, \cdots$ 都是 $\widetilde{P}$ 的暂留状态, 所以 $\widetilde{\Pi} e_i = 0(i \geqslant 1)$, 从而由 $\widetilde{P}$ 之无耗损性得

$$\mathbf{1} = \widetilde{\Pi}\mathbf{1} = \widetilde{\Pi} \sum_{i \in E} e_i = \widetilde{\Pi} e_0 = \widetilde{y} = \begin{pmatrix} 1 \\ f_{0,1}^* \\ f_{0,2}^* \\ \vdots \end{pmatrix}.$$

所以 $f_{0,i}^* = 1(i \geqslant 1)$, 再用推论 5.2 得 $f_{0,0}^* = 1$, 此即 0 是常返状态.

**附注 5.1** 本章所讨论的问题都是从 $P$ 链的概率空间 $(\Omega^*, \mathscr{F}^*, P^i)$ 出发的, 实际上是从转移矩阵 $P = (p_{i,j}, i, j \in E)$ 出发的. 所以本章的结论对一般的时齐的具有转移矩阵 $P$ 的 Markov 链来说, 也是成立的.

## §6 习题及应用

1. 证明命题 2.2.

2. 证明: 任给 $P$ 链 $(\Omega^*, \mathscr{F}^*, X_n, \theta_m, P^i, i \in E, n, m \geqslant 0), \{X_n, n \geqslant 0\}$ 是概率空间 $(\Omega^*, \mathscr{F}^*, P^i)$ 上的时齐的以 $E$ 为状态空间的具有转移矩阵 $P$ 的以 $(\delta_{i,j}, j \in E)$ 为初始分布的 Markov 链.

本节以下剩余部分恒设 $E$ 是可数集, $P = (p_{i,j}, i, j \in E)$ 是 $E \times E$ 上的转移矩阵, $P^n = P^{(n)} = (p_{i,j}^{(n)}, i, j \in E)$ 是 $n$ 步转移矩阵, $\Pi = \lim\limits_{n \to \infty} \dfrac{1}{n} \sum\limits_{k=1}^{n} P^k, E$ 分解成: $E = N \bigcup R = N \bigcup (R^0 \bigcup R^+) = N \bigcup \left( \bigcup\limits_m R_m^0 \right) \bigcup (R_n^+), N, R, R^0, R^+, R_m^0, R_n^+$ 的定义如 §4.

3. 若 $E$ 是有限集, 则 $R^0 = \varnothing, R^+ \neq \varnothing$.

4. 若 $E$ 是有限集, 则下列陈述等价:

(1) $\Pi$ 没有 0 列;

(2) $E$ 不含暂留状态.

5. 若 $E$ 是有限集, 则下列陈述等价:

(1) $\Pi$ 行行一样, 即 $\Pi = \mathbf{1}\pi', \pi'$ 是行向量;

(2) $E$ 恰含一个正类 ($N$ 可能非空).

6. 设 $E$ 为有限集, 则下列陈述等价:

(1) $\Pi > 0$;

(2) 对任何 $i, j \in E$, 存在正整数 $n$ (可以依赖于 $i, j$), 使 $p_{i,j}^{(n)} > 0$;

(3) $P$ 是不可约的;

(4) $E = R_1^+$;

(5) $\Pi > 0$ 且 $\Pi = \mathbf{1}\pi'$ 行行一样.

7. 若 $E$ 是有限集, 则下列陈述等价:

(1) $P$ 是不可约的无周期的;

(2) 存在正整数 $M$ 使 $P^M > 0$.

8. 若 $E$ 恰由 $S$ 个元素所构成, 则有:

(1) $f^*_{i,i} > 0 \Rightarrow$ 存在正整数 $m \leqslant S$, 使 $p^{(m)}_{i,i} > 0$;

(2) $f^*_{i,j} > 0$ 且 $i \neq j \Rightarrow$ 存在正整数 $m < S$, 使 $p^{(m)}_{i,j} > 0$.

9. 若 $E$ 是有限集, 则 1 是 $P$ 的特征根, 且对 $P$ 的任何特征根 $\lambda$, 都有 $|\lambda| \leqslant 1$.

10. 若 $E$ 是有限集, 则

$$\Pi = \lim_{\lambda \uparrow 1}(1-\lambda)(I-\lambda P)^{-1},$$

其中 $I$ 是单位矩阵.

11. 整数格子点上的一步自由随机徘徊. 设 $E = \{0, \pm 1, \pm 2, \cdots\}, P = (p_{i,j}, i,j \in E), p_{i,i-1} = q_i, p_{i,i+1} = p_i, p_i > 0, q_i > 0, p_i + q_i = 1 (\forall i \in E)$. 当 $|i-j| > 1$ 或 $i = j$ 时, $p_{i,j} = 0$.

试证:

(1) $P$ 是不可约的;

(2) $P$ 的周期 $d = 2$.

此外, 请找出 $E = R_1^+$ (或 $E = R_1^0$ 或 $E = N$) 的充分必要条件.

12. 在 0 处具有反射壁的随机徘徊. 设 $E = \{0, 1, 2, \cdots\}, P = (p_{i,j}, i, j \in E), p_{0,0} = a_0, p_{0,1} = b_0, p_{0,j} = 0$ (当 $j > 1$ 时). 当 $i \geqslant 1$ 时有: $p_{i,i-1} = a_i, p_{i,i+1} = b_i, p_{i,j} = 0$ (当 $i = j$ 或 $|i-j| > 1$), $a_i > 0$ (当 $i \geqslant 1$), $b_i > 0$ (当 $i \geqslant 0$), $a_i + b_i = 1$ (对一切 $i \in E$).

(1) 证明 $P$ 是不可约的;

(2) 找出 $E = N$ (或 $E = R_1^0$ 或 $E = R_1^+$) 的充分必要条件;

(3) 求出 $\Pi$.

13. 在 0 处具有吸收壁的随机徘徊. 设 $E = \{0, 1, 2, \cdots\}, P = (p_{i,j}, i, j \in E), p_{0,0} = 1, p_{0,j} = 0$, (当 $j \geqslant 1$ 时), 当 $i \geqslant 1$ 时, $p_{i,i-1} = q_i, p_{i,i+1} = p_i, p_{i,j} = 0$ (当 $|i-j| > 1$ 或 $i = j$), $p_i > 0, q_i > 0, p_i + q_i = 1$.

(1) 试证 $P$ 是可约的, 0 是正状态, 其他状态都是暂留状态;

(2) 求出 $\Pi$.

14. 更新过程 (离散时间的). 考虑一件 "设备" (它可能是一台机器, 也可能是一个零件, 或一个细胞), 当它使用了 $i$ "年" 以后, 在下一年内坏掉的概率为 $q_i$, 不坏的概率为 $p_i$. 而一旦坏掉, 就立刻换上一个同类的设备. 假设最初安

装上去的是一个新设备. 令 $X_n$ 是第 $n$ 年时在用的设备的年龄, $(n \geqslant 0), p_i > 0, q_i > 0, p_i + q_i = 1$ (对一切 $i = 0, 1, 2, \cdots$). 试证 $\{X_n, n \geqslant 0\}$ 是一个时齐的以 $E = \{0, 1, 2, \cdots\}$ 为状态空间, 以 $P = (p_{i,j}, i, j \in E)$ 为转移矩阵的 Markov 链, 其中 $p_{i,i+1} = p_i, p_{i,0} = q_i (\forall i \in E)$. 此外, 找出 $E = N$ (或 $E = R_1^0$ 或 $R_1^+$) 的充分必要条件并算出 $\Pi$.

15. 考虑一个服务机构 (可能是一个售货柜台, 也可能是一个港口码头, ……). 每个来到此服务机构的 "顾客" (可能是一个人, 也可能是一条船, ……) 要求服务的时间是固定长度 $l$ (例如每条船的装卸货时间). 再设在长为 $l$ 的时间内新来到的顾客数 $Y$ 是一个以 $\{k_0, k_1, k_2, \cdots\}$ 为概率分布的随机变量, 其中 $k_0 > 0, k_2 + k_3 + \cdots > 0, \sum\limits_{i=0}^{\infty} k_i = 1$. 记 $X_n$ 为第 $n$ 个到来的顾客服务完后在此服务机构等待服务的顾客数 (即队的长度), $n \geqslant 1$. 试证: $\{X_n, n \geqslant 1\}$ 是一个时齐的以 $E = \{0, 1, 2, \cdots\}$ 为状态空间的以

$$P = \begin{pmatrix} k_0 & k_1 & k_2 & k_3 & \cdots \\ k_0 & k_1 & k_2 & k_3 & \cdots \\ 0 & k_0 & k_1 & k_2 & \cdots \\ 0 & 0 & k_0 & k_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

为转移矩阵的 Markov 链, 且 $P$ 是不可约的. 此外, 试再求出 $\Pi$.

# *第七章　可数状态的 Markov 过程

## §1　转移矩阵的连续性及可微性

在这一章中, 恒设时间参数集 $\boldsymbol{T}=[0,\infty)$, 状态空间 $E$ 是可数无穷集或有限集. $f'(t)$ 表 $f(t)$ 的微商. $I$ 是单位矩阵, $\delta_{i,j}=0$ 或 1 视 $i\neq j$ 或 $i=j$ 而定.

**定义 1.1**　对任何 $i,j\in E$, 设

$$p_{i,j}(\cdot):\boldsymbol{T}\mapsto[0,1].$$

称 $E\times E$ 上的函数项矩阵 $P(t)=(p_{i,j}(t),i,j\in E)$ 是**准转移矩阵**, 如果

(1) $\sum\limits_{j\in E}p_{i,j}(t)\leqslant 1,(\forall i\in E,t\in\boldsymbol{T})$;

(2) $P(s)P(t)=P(s+t)(\forall s,t\in\boldsymbol{T})$.

特别地, 若 (1) 中的 $\leqslant$ 代之以 $=$, 则称 $P(t)$ 是**转移矩阵**. 称 (2) 为 Kolmogorov-Chapmann (K − C) 方程式.

注意: 第六章所言之转移矩阵的元素是常数, 而第七章之转移矩阵的元素都是函数, 由于不会混淆, 故为了简单起见, 此二名词不加区别.

本章凡遇矩阵之间比较大小, 极限、连续微商、积分 $\cdots$ 等等, 都是逐元意义下的. 例如对于 $Q(t)=(q_{i,j}(t),i,j\in E),P(t)=(p_{i,j}(t),i,j\in E),Q(t)\geqslant P(t)$ 是指 $q_{i,j}(t)\geqslant p_{i,j}(t)(\forall i,j\in E)$. 本章恒令 $I$ 为单位矩阵, $\mathbf{1}=\begin{pmatrix}1\\1\\\vdots\end{pmatrix}$ 为分量恒为 1 的列向量, 其维数视当时的具体情况而定.

**定义 1.2** 设 $\{X_t, t \in \boldsymbol{T}\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的取值于 $E$ 的一族随称变量, 如果对任意的 $n \geqslant 2, 0 \leqslant t_1 < t_2 < \cdots < t_n$, 及任意的 $\{i_1, \cdots, i_n\} \subset E$, 都有

$$\begin{aligned} & P(X_{t_n} = i_n | X_1 = i_1, \cdots, X_{t_{n-1}} = i_{n-1}) \\ = & P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}) \end{aligned} \tag{1.1}$$

(当 (1.1) 左方有意义时), 则称$\{X_t, t \in \boldsymbol{T}\}$是**一个连续时间的可数状态的 Markov 过程**, 简称 **Markov 过程**, 称 $E$ 为其状态空间.

特别地, 如果存在转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$ 使

$$P(X_{s+t} = j | X_s = i) = p_{i,j}(t) \quad (i, j \in E, s, t \in \boldsymbol{T}), \tag{1.2}$$

(当等式左方有意义时), 则称 $\{X_t, t \in \boldsymbol{T}\}$ 是一个**具有时齐的转移概率的可数状态的 Markov 过程**, 简称**时齐的可数状态的 Markov 过程**. 本章只研究这类 Markov 过程 $\{X_t, t \in \boldsymbol{T}\}$, 称 $P(t)$ 为它的转移概率矩阵. $\{P(X_0 = i), i \in E\}$ 为**其初始分布**.

正如第六章 §1 末尾所指出的, 任意给定一个概率分布 $\{\mu_i, i \in E\}$ 和一个转移矩阵 $P(t)$, 恒存在一个概率空间 $(\Omega, \mathscr{F}, P)$ 以及定义在其上的时齐的可数状态的 Markov 过程 $\{X_t, t \in \boldsymbol{T}\}$, 它的转移概率矩阵为 $P(t)$, 它的初始分布为 $\{\mu_i, i \in E\}$.

与第六章 §2 类似, 任意给定一个转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$, 可以构造一个 $P(t)$ 过程及与其相应的 $P(t)$ 过程概率空间 $(\Omega, \mathscr{F}, P^i)(i \in E)$ 如下:

令 $\Omega = E^{\boldsymbol{T}}, \Omega$ 中每一个元素 $\omega$ 都是由 $\boldsymbol{T}$ 到 $E$ 的映射. 任取 $t \in \boldsymbol{T}, A \subset E$, 记

$$\begin{bmatrix} t \\ A \end{bmatrix} = \{\omega \in \Omega : \omega(t) \in A\},$$

特别地, 当 $A = \{j\}$ 是 $E$ 中的单点集时, 则记

$$\begin{bmatrix} t \\ A \end{bmatrix} = \begin{bmatrix} t \\ j \end{bmatrix}.$$

若 $0 \leqslant t_1 < t_2 < \cdots < t_n, A_i \subset E(i = 1, \cdots, n)$, 则记

$$\begin{bmatrix} t_1, \cdots, t_n \\ A_1, \cdots, A_n \end{bmatrix} = \bigcap_{k=1}^{n} \begin{bmatrix} t_k \\ A_k \end{bmatrix}.$$

令 $\mathscr{F}=\sigma\left(\begin{bmatrix}t\\j\end{bmatrix},t\in\boldsymbol{T},j\in E\right)$, 再令

$$P^i\left(\begin{bmatrix}t_1,\cdots,t_n\\i_1,\cdots,i_n\end{bmatrix}\right)=p_{i,i_1}(t_1)p_{i_1,i_2}(t_2-t_1)\cdots p_{i_{n-1},i_n}(t_n-t_{n-1})$$

($\{i,i_1,\cdots,i_n\}\subset E,\{t_1,\cdots,t_n\}\subset\boldsymbol{T}$), 则用文献 [35] 第四章定理 3.1 的 Kolmogorov 定理, $P^i$ 可唯一地扩张到 $\mathscr{F}$ 去而成一概率测度, 此扩张之概率测度仍用 $P^i$ 记之. 于是得概率空间 $(\Omega,\mathscr{F},P^i)(i\in E)$.

再定义一族由 $\Omega$ 到 $\Omega$ 的推移算子 $\theta_t$ 如下: $(\theta_t\omega)(s)=\omega(s+t)(s,t\in\boldsymbol{T})$. 令 $X_t$ 为 $\Omega$ 到 $\Omega$ 的坐标变换: $X_t(\omega)=\omega(t)(\omega\in\Omega,t\in\boldsymbol{T})$. 称 $(\Omega,\mathscr{F},X_t,\theta_t,P^i)(i\in E)$ 为 $P(t)$ 过程, $(\Omega,\mathscr{F},P^i)$ 为 $P(t)$ 过程的概率空间.

类似于第六章命题 2.1, 有

**命题 1.1** 对 $P(t)$ 过程的概率空间 $(\Omega,\mathscr{F},P^i)$, 有下列 Markov 性:

$$(M):P^i\left(A\bigcap\begin{bmatrix}s\\j\end{bmatrix}\bigcap\theta_t^{-1}B\right)=P^i\left(A\bigcap\begin{bmatrix}s\\j\end{bmatrix}\right)P^j(\theta_{t-s}^{-1}B),\tag{1.3}$$

$$\left(i,j\in E,t\geqslant s\geqslant 0,B\in\mathscr{F},A\in\mathscr{F}_s\overset{\text{def.}}{=\!=}\sigma\left(\begin{bmatrix}u\\j\end{bmatrix},j\in E,u\leqslant s\right).\right)$$

**证** (1) 先取 $A=\begin{bmatrix}u\\k\end{bmatrix},u\leqslant s,k\in E$ 固定. 对任何 $B=\begin{bmatrix}v\\l\end{bmatrix},v\geqslant 0,l\in E$, 有

$$\begin{aligned}P^i\left(A\bigcap\begin{bmatrix}s\\j\end{bmatrix}\bigcap\theta_t^{-1}B\right)&=P^i\left(\begin{bmatrix}u,&s,&t+v\\k,&j,&l\end{bmatrix}\right)\\&=p_{i,k}(u)p_{k,j}(s-u)p_{j,l}(t+v-s)\\&=P^i\left(A\bigcap\begin{bmatrix}s\\j\end{bmatrix}\right)P^j(\theta_{t-s}^{-1}B).\end{aligned}$$

而使 (1.3) 成立之 $B$ 构成一个 $\sigma$ 代数, 所以对任何 $A=\begin{bmatrix}u\\k\end{bmatrix},u\leqslant s,k\in E,B\in\mathscr{F}$, (1.3) 都成立.

(2) 又因为对固定的 $B\in\mathscr{F}$, 使 (1.3) 成立的 $A$ 亦构成一个 $\sigma$ 代数, 所以由 (1) 得知: 对任何 $A\in\mathscr{F}_s,B\in\mathscr{F}$, (1.3) 都成立. 命题证毕.

特别地, 取 $s=t=t_n$,

$$A=\begin{bmatrix} t_1, \cdots, t_{n-1} \\ A_1, \cdots, A_{n-1} \end{bmatrix},$$

$$B=\begin{bmatrix} t_{n+1}-t_n, \cdots, t_{n+m}-t_n \\ A_{n+1}, \quad \cdots, \quad A_{n+m} \end{bmatrix},$$

$$0 \leqslant t_1 < t_2 < \cdots < t_n < t_{n+1} < \cdots < t_{n+m},$$

$$A_k \subset E, \quad (k=1,2,\cdots,n+m),$$

则 (1.3) 化为

$$(M_1): P^i\left(\begin{bmatrix} t_1, \cdots, t_{n-1}, t_n, t_{n+1}, \cdots, t_{n+m} \\ A_1, \cdots, A_{n-1}, j, A_{n+1}, \cdots, A_{n+m} \end{bmatrix}\right)$$
$$= P^i\left(\begin{bmatrix} t_1, \cdots, t_{n-1}, t_n \\ A_1, \cdots, A_{n-1}, j \end{bmatrix}\right) P^j\left(\begin{bmatrix} t_{n+1}-t_n, \cdots, t_{n+m}-t_n \\ A_{n+1}, \quad \cdots, \quad A_{n+m} \end{bmatrix}\right).$$

思考题: (1.1) 和 (1.3) 是 Markov 性的不同表述, 它们之间有何关系? 时齐的可数状态的 Markov 过程与 $P(t)$ 过程 $(\Omega,\mathscr{F},X_t,\theta_t,P^i)(i\in E)$ 有何关系?

根据 $P(t)$ 过程和时齐的可数状态的以 $P(t)$ 为转移概率矩阵的 Markov 的定义, 它们的概率规律本质上都是由转移矩阵 $P(t)$ 决定的 (除了不起本质作用的初始分布 $\{\mu_i, i\in E\}$ 以外), 所以, 我们研究 $P(t)$ 的性质, 特别是它作为 $t$ 的函数的分析性质, 是十分重要的. 这一节, 我们主要是研究它的连续性与可微性.

**定义 1.3** 称准转移矩阵 $P(t)=(p_{i,j}(t),i,j\in E)$ 是**标准的**, 如果 $P(0+)=P(0)=I$, 即是

$$\lim_{t\to 0+} p_{i,j}(t)=p_{i,j}(0)=\delta_{i,j}, \quad (\forall i,j\in E). \tag{1.4}$$

**定理 1.1** 设 $P(t)=(p_{i,j}(t),i,j\in E)$ 是标准的转移矩阵, 则对任何 $i\in E, J\subset E, 0\leqslant s,t<\infty$, 都有

$$|p_{i,J}(s)-p_{i,J}(t)| \leqslant 1-p_{i,i}(|t-s|),$$

其中

$$p_{i,J}(t)=\sum_{j\in J} p_{i,j}(t).$$

因此, $p_{i,J}(t)$ 对 $t$ 来说在 $[0,\infty)$ 一致连续 (当然在 0 点只是右连续, 因为 $p_{i,J}(\cdot)$ 在 $(-\infty,0)$ 内无定义) 而且对 $J\subset E$ 来说是等度连续. 特别地, 对任何 $i,j\in E, p_{i,j}(t)$ 对 $t$ 来说在 $[0,\infty)$ 一致连续, 而且对 $j\in E$ 来说是等度连续.

**证** 不失普遍性可设 $0 \leqslant s < t$. 由 (K − C) 方程式有

$$p_{i,J}(t) \geqslant p_{i,i}(t-s)p_{i,J}(s),$$

所以

$$\begin{aligned} p_{i,J}(t) - p_{i,J}(s) &\geqslant (p_{i,i}(t-s) - 1)p_{i,J}(s) \\ &\geqslant p_{i,i}(t-s) - 1 \quad (\forall J \subset E). \end{aligned}$$

由于上式对任何 $J \subset E$ 皆成立, 特别地, 对 $J^c = E - J$ 也成立, 此即

$$p_{i,J^c}(t) - p_{i,J^c}(s) \geqslant p_{i,i}(t-s) - 1.$$

综合上述两个不等式即得定理 1.1 的结论.

**定理 1.2** 对任何标准转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$ 都有
(1) $\lim\limits_{t\to\infty} P(t) = \Pi = (\pi_{i,j}, i, j \in E)$ 存在;
(2) $\Pi P(t) = P(t)\Pi = \Pi^2 = \Pi$.

**证** (1) 由 $P(t)$ 满足 (K − C) 方程式知:

$$p_{i,i}(t) \geqslant p_{i,i}\left(\frac{t}{2}\right)^2 \geqslant \cdots \geqslant p_{i,i}\left(\frac{t}{2^n}\right)^{2^n}$$

$(\forall i \in E, t > 0)$. 再由 $P(t)$ 的标准性知

$$\lim_{t\to 0+} p_{i,i}(t) = 1 \quad (\forall i \in E).$$

所以

$$p_{i,i}(t) > 0 \quad (\forall t > 0, i \in E).$$

今固定任意一对 $\{i,j\} \subset E$, 记 $f(t) = p_{i,j}(t)$, 要证 $\lim\limits_{t\to\infty} f(t)$ 存在. 为此只需证明

$$\lim_{t_n \uparrow \infty} f(t_n) \quad 存在.$$

事实上, 固定任一 $\tau > 0$, 由 $p_{i,i}(\tau) > 0 (\forall i \in E)$ 及第六章定理 5.4 知

$$\lim_{n\to\infty} P(n\tau) = \lim_{n\to\infty} P(\tau)^n \quad 存在.$$

而由定理 1.1 知 $f(t)$ 在 $t \in [0, \infty)$ 上一致连续. 所以任意给定 $\varepsilon > 0$, 存在 $\tau = \tau(\varepsilon) > 0$ 使

$$|f(t) - f(s)| \leqslant \varepsilon, \quad (当\ |t-s| \leqslant \tau).$$

对每个 $n$, 可选非负整数 $k_n$, 使

$$k_n\tau \leqslant t_n < (k_n+1)\tau.$$

所以

$$f(k_n\tau)-\varepsilon \leqslant f(t_n) \leqslant f(k_n\tau)+\varepsilon \quad (n=1,2,\cdots).$$

但是

$$\lim_{n\to\infty} f(k_n\tau) \quad \text{存在, 记此极限为 } l.$$

故

$$l-\varepsilon \leqslant \liminf_{n\to\infty} f(t_n) \leqslant \limsup_{n\to\infty} f(t_n) \leqslant l+\varepsilon.$$

由 $\varepsilon>0$ 可任意小得知 $\lim\limits_{t_n\uparrow\infty} f(t_n)$ 存在.

(2) 由 (1) 及 (K − C) 方程式 $P(s+t)=P(t)P(s)$ 并用控制收敛定理可得

$$\begin{aligned}\varPi &= \lim_{s\to\infty} P(s+t) = \lim_{s\to\infty} P(t)P(s)\\ &= P(t)\left(\lim_{s\to\infty} P(s)\right) = P(t)\varPi.\end{aligned}$$

仿上, 在 $P(s+t)=P(s)P(t)$ 中令 $s\to\infty$ 并利用 Fatou 引理可得

$$\begin{aligned}\varPi &= \lim_{s\to\infty} (P(s)P(t)) \geqslant \left(\lim_{s\to\infty} P(s)\right) P(t)\\ &= \varPi P(t) \quad (0<t<\infty).\end{aligned}$$

但是 $(\varPi P(t))\mathbf{1} = \varPi(P(t)\mathbf{1}) = \varPi\mathbf{1}$, 所以

$$\varPi = \varPi P(t) \quad (0<t<\infty).$$

上式对 $t\to\infty$ 取极限并利用控制收敛定理即得 $\varPi=\varPi^2$. 定理证毕.

下面我们要研究 $P(t)$ 的可微性. 先证一个引理.

**引理 1.1** 设函数 $f:[0,\infty)\mapsto[-\infty,\infty]$ 满足条件:

(1) $f(u+v)\leqslant f(u)+f(v) \quad (\forall u,v\geqslant 0)$;

(2) $\lim\limits_{t\to 0+} f(t)=0$,

则

$$\lim_{t\to 0+}\frac{f(t)}{t} = \sup_{t\geqslant 0}\frac{f(t)}{t}. \tag{1.5}$$

**证** 令 $\eta(u)=\sup\limits_{0\leqslant t\leqslant u} f(t)$. 任取 $t>0$, 则对任何 $u\in(0,t)$, 均存在正整数 $n$, 使

$$nu < t \leqslant (n+1)u. \tag{1.6}$$

反复利用条件 (1) 及 (1.6) 式得

$$f(t) \leqslant f(t-nu)+f(nu) \leqslant nf(u)+\eta(u),$$

所以

$$\frac{f(t)}{t} \leqslant \frac{\eta(u)}{t}+\frac{n}{t}f(u). \tag{1.7}$$

令 $u \to 0+$ 得

$$\frac{f(t)}{t} \leqslant \liminf_{u\to 0+}\frac{f(u)}{u}.$$

所以

$$\sup_{t\geqslant 0}\frac{f(t)}{t} \leqslant \liminf_{u\to 0+}\frac{f(u)}{u},$$

从而

$$\lim_{t\to 0+}\frac{f(t)}{t}=\sup_{t>0}\frac{f(t)}{t}.$$

**定理 1.3** 对任何标准的转移矩阵 $P(t)=(p_{i,j}(t), i,j\in E)$, 恒有

$$\lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t}=q_i \quad 存在 \quad (\forall i\in E), \tag{1.8}$$

(此处 $q_i\in[0,\infty]$) 且 $p_{i,i}(t)\geqslant e^{-q_it}$.

**证** 由 $P(t)$ 的标准性及 $P(t)=P\left(\frac{t}{2^n}\right)^{2^n}$ 得知

$$p_{i,i}(t)>0 \quad (\forall i\in E, t\geqslant 0).$$

所以固定任意 $i\in E$, 可令 $f(t)=-\log p_{i,i}(t)$. 由于 $\lim\limits_{t\to 0+}p_{i,i}(t)=1$ 及 $p_{i,i}(s+t)\geqslant p_{i,i}(s)p_{i,i}(t)$ 可知 $f(t)$ 满足引理 1.1 中的条件, 所以

$$\lim_{t\to 0+}\frac{f(t)}{t}=\sup_{t>0}\frac{f(t)}{t}.$$

令

$$q_i=\sup_{t>0}\frac{f(t)}{t},$$

则

$$f(t)=q_it+0(t) \quad (当\ t\to 0+时),$$

且

$$f(t)\leqslant q_it \quad (\forall t>0).$$

所以

$$p_{i,i}(t)=e^{-q_it+o(t)}=e^{-q_it}+o(t),$$

从而

$$\frac{1-p_{i,i}(t)}{t}=\frac{1-e^{-q_it}}{t}+o(1)\quad(t\to 0+).$$

所以

$$\lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t}=q_i\quad 存在,$$

$q_i\in[0,\infty]$.

由 $f(t)\leqslant q_it$ 即得 $p_{i,i}(t)\geqslant e^{-q_it}(t\geqslant 0)$.

**推论 1.1** 若 $q_i=0$, 则 $p_{i,i}(t)\equiv 1$.

**引理 1.2** 设 $P(t)=(p_{i,j}(t),i,j\in E)$ 是一个标准的转移矩阵, 令

$$\mathscr{G}=\left\{J\subset E:\lim_{t\to 0+}\sup_{j\in J}(1-p_{j,j}(t))=0\right\},$$
$$p_{i,J}(t)=\sum_{j\in J}p_{i,j}(t)\quad(J\subset E).$$

任给 $J\in\mathscr{G},\varepsilon<\dfrac{1}{4}$, 如果 $\tau>0$ 满足

$$1-p_{j,j}(t)<\varepsilon\quad(\forall 0\leqslant t\leqslant\tau,j\in J),$$

则对一切 $0<v\leqslant\tau,0<\dfrac{u}{v}\leqslant\varepsilon,K\subset J,i\in J-K$, 有

$$(1-4\varepsilon)\frac{p_{i,K}(u)}{u}\leqslant\frac{p_{i,K}(v)}{v}.\tag{1.9}$$

**证** 任取 $i\in E,G\subset E$, 令

$$f_1(i,G)=p_{i,G}(u),\quad f_{m+1}(i,G)=\sum_{j\in K}f_m(i,\{j\})p_{j,G}(u),$$

则对任意固定的 $i\in E,m\geqslant 1,f_m(i,\cdot)$ 有完全可加性.

下面我们对 $m$ 作归纳法来证明

$$\begin{aligned}p_{i,G}(t)=&\sum_{l=1}^{m}\sum_{j\in K}f_l(i,\{j\})p_{j,G}(t-lu)\\&+\sum_{j\overline{\in}K}f_m(i,\{j\})p_{j,G}(t-mu),\end{aligned}\tag{1.10}$$

其中 $m\geqslant 1,t\geqslant mu,G\subset E$.

当 $m=1$ 时, (1.10) 显然成立. 设 (1.10) 对 $m$ 成立, 要证 (1.10) 对 $m+1$ 也成立. 事实上

$$\begin{aligned}
&\sum_{l=1}^{m+1}\sum_{j\in K} f_l(i,\{j\})p_{j,G}(t-lu)+\sum_{j\in K} f_{m+1}(i,\{j\})p_{j,G}(t-(m+1)u)\\
=&\sum_{l=1}^{m+1}\sum_{j\in K} f_l(i,\{j\})p_{j,G}(t-lu)+\sum_{j\in K} f_{m+1}(i,\{j\})p_{j,G}(t-(m+1)u)\\
&+\sum_{j\in K} f_m(i,\{j\})p_{j,G}(t-mu)-\sum_{j\in E}\sum_{k\in K} f_m(i,\{k\})p_{k,j}(u)p_{j,G}(t-(m+1)u)\\
=&\sum_{l=1}^{m+1}\sum_{j\in K} f_l(i,\{j\})p_{j,G}(t-lu)+\sum_{j\in K} f_{m+1}(i,\{j\})p_{j,G}(t-(m+1)u)\\
&+\sum_{j\in K} f_m(i,\{j\})p_{j,G}(t-mu)-\sum_{j\in E} f_{m+1}(i,\{j\})p_{j,G}(t-(m+1)u)\\
=&\sum_{l=1}^{m}\sum_{j\in K} f_l(i,\{j\})p_{j,G}(t-lu)+\sum_{j\in K} f_m(i,\{j\})p_{j,G}(t-mu)\\
=&\,p_{i,G}(t).
\end{aligned}$$

归纳法完成.

现在, 我们反复用 (1.10) 来证明引理 1.2.

(1) 在 (1.10) 中令 $G=K, t=v, m=n_0$, $\left(n_0=\left[\dfrac{u}{v}\right]\right.$ 是 $\dfrac{u}{v}$ 的最大整数部分,$\Big)$ 则得

$$\begin{aligned}
p_{i,K}(v)&\geqslant\sum_{l=1}^{n_0}\sum_{j\in K} f_l(i,\{j\})p_{j,K}(v-lu)\\
&\geqslant\sum_{l=1}^{n_0}\sum_{j\in K} f_l(i,\{j\})p_{j,j}(v-lu)\\
&\geqslant\sum_{l=1}^{n_0} f_l(i,K)\quad(1-\varepsilon).
\end{aligned}\tag{1.11}$$

所以若注意 $i\in J-K, J\in\mathscr{G}$, 则由 (1.11) 得

$$\sum_{l=1}^{n_0} f_l(i,K)\leqslant\frac{p_{i,K}(v)}{1-\varepsilon}\leqslant\frac{\varepsilon}{1-\varepsilon}.\tag{1.12}$$

(2) 在 (1.10) 中令 $G=\{i\}, t=mu (1\leqslant m\leqslant n_0-1)$, 得

$$p_{i,\{i\}}(mu)\leqslant\sum_{l=1}^{n_0-1} f_l(i,K)+f_m(i,\{i\}).\tag{1.13}$$

但是由 $1 \leqslant m \leqslant n_0+1$ 得

$$p_{i,\{i\}}(mu) > 1-\varepsilon. \tag{1.14}$$

由 (1.12)、(1.13) 和 (1.14) 得

$$f_m(i,\{i\}) > 1-\varepsilon-\frac{\varepsilon}{1-\varepsilon} > \frac{1-3\varepsilon}{1-\varepsilon} \quad (1 \leqslant m \leqslant n_0-1). \tag{1.15}$$

(3) 在 (1.10) 中令 $G=K, m=n_0, t=v$ 并注意 $i \in J-K$ 及 $f_m$ 之定义, 得

$$\begin{aligned}
p_{i,K}(v) &\geqslant \sum_{l=1}^{n_0}\sum_{j\in K} f_l(i,\{j\})p_{j,K}(v-lu) \\
&\geqslant (1-\varepsilon)\sum_{j\in K} f_1(i,\{j\}) + \sum_{l=2}^{n_0}\sum_{j\in K} f_l(i,\{j\})p_{j,K}(v-lu) \\
&= (1-\varepsilon)p_{i,K}(u) + \sum_{l=2}^{n_0}\sum_{j\in K} f_l(i,\{j\})p_{j,K}(v-lu) \\
&\geqslant (1-\varepsilon)p_{i,K}(u) + \sum_{l=2}^{n_0}\sum_{j\in K} f_{l-1}(i,\{i\})p_{i,j}(u)p_{j,K}(v-lu).
\end{aligned} \tag{1.16}$$

因此由 (1.15)、(1.16) 及引理的假设得

$$\begin{aligned}
p_{i,K}(v) &\geqslant (1-\varepsilon)p_{i,K}(u) + \sum_{l=2}^{n_0}\frac{1-3\varepsilon}{1-\varepsilon}(1-\varepsilon)\sum_{j\in K} p_{i,j}(u) \\
&= (1-\varepsilon)p_{i,K}(u) + (n_0-1)(1-3\varepsilon)p_{i,K}(u),
\end{aligned}$$

所以

$$\frac{p_{i,k}(v)}{v} \geqslant (1-3\varepsilon)\frac{n_0 u}{v}\cdot\frac{p_{i,K}(u)}{u}.$$

但是

$$\frac{n_0 u}{v} = \left[\frac{v}{u}\right]\frac{u}{v} \geqslant 1-\frac{u}{v} \geqslant 1-\varepsilon,$$

所以

$$(1-4\varepsilon)\frac{p_{i,K}(u)}{u} \leqslant \frac{p_{i,K}(v)}{v}.$$

引理 1.2 证毕.

**定理 1.4** 设 $P(t)=(p_{i,j}(t), i,j\in E)$ 是一个标准的转移矩阵, $\mathscr{G}$ 和 $p_{i,J}$ 之定义如引理 1.2, 则对任意的 $i \overline{\in} J \in \mathscr{G}$, 恒有

(1) 当 $K \subset J$ 时,

$$\lim_{t\to 0+}\frac{p_{i,K}(t)}{t} = q_{i,K} \quad 存在, \tag{1.17}$$

而且 $0 \leqslant q_{i,K} < \infty$, (1.17) 中之极限对 $K \subset J$ 一致成立;

(2) $q_{i,K}$ 对 $K \subset J$ 来说, 有完全可加性.

**证** (1) 令 $G=\{i\}\bigcup J$, 则 $G\in\mathscr{G}$, 所以对任何 $\varepsilon<\dfrac{1}{4}$, 必存在 $\tau=\tau(\varepsilon,G)$, 使

$$1-p_{j,j}(t)<\varepsilon\quad(\forall j\in G,0\leqslant t\leqslant\tau).$$

所以由引理 1.2 知

$$\frac{(1-4\varepsilon)p_{i,K}(u)}{u}\leqslant\frac{p_{i,K}(v)}{v}\quad\left(\text{当 }0<v\leqslant\tau,0<\frac{u}{v}\leqslant\varepsilon\right).$$

在上式中令 $u\to0+$, 取上极限即得

$$(1-4\varepsilon)\limsup_{u\to0+}\frac{p_{i,K}(u)}{u}\leqslant\frac{p_{i,K}(v)}{v}\quad(\text{当 }0<v\leqslant\tau).\tag{1.18}$$

再把 (1.18) 对 $v\to0+$ 取下极限, 并注意 $\varepsilon>0$ 可以任意小得知

$$\lim_{t\to0+}\frac{p_{i,K}(t)}{t}=q_{i,K}\quad\text{存在},$$

且 $0\leqslant q_{i,K}<\infty$. 且 (1.18) 化为

$$\frac{p_{i,K}(v)}{v}\geqslant(1-4\varepsilon)q_{i,K}\quad(\text{当 }0<v\leqslant\tau).\tag{1.19}$$

显然 $q_{i,K}\leqslant q_{i,J}$, 所以

$$\frac{p_{i,K}(v)}{v}-q_{i,K}\geqslant-4\varepsilon q_{i,J}\quad(\text{当 }0<v\leqslant\tau).\tag{1.20}$$

但 $K$ 可为 $J$ 的任意子集, 所以在 (1.20) 中以 $J-K$ 代 $K$ 得

$$\frac{p_{i,J}(v)-p_{i,K}(v)}{v}-(q_{i,J}-q_{i,K})\geqslant-4\varepsilon q_{i,J},$$

亦即

$$\frac{p_{i,K}(v)}{v}-q_{i,K}\leqslant\left(\frac{p_{i,J}(v)}{v}-q_{i,J}\right)+4\varepsilon q_{i,J}\quad(\text{当 }0<v\leqslant\tau).\tag{1.21}$$

由 (1.20) 和 (1.21) 得

$$\left|\frac{p_{i,K}(v)}{v}-q_{i,K}\right|\leqslant4\varepsilon q_{i,J}+\left|\frac{p_{i,J}(v)}{v}-q_{i,J}\right|\quad(\text{当 }0<v\leqslant\tau).\tag{1.22}$$

(1.22) 说明了

$$\lim_{v\to0+}\frac{p_{i,K}(v)}{v}=q_{i,K}\quad(\text{对 }K\subset J\text{ 一致成立}).$$

(2) 显然, $q_{i,K}$ 对 $K\subset J$ 来说具有有限可加性. 若再注意 (1.19) 则对任何 $K_n\subset J,K_n\downarrow\varnothing$ 有

$$\limsup_{n\to\infty}q_{i,K_n}\leqslant\limsup_{n\to\infty}\frac{1}{1-\varepsilon}\frac{p_{i,K_n}(v)}{v}=0,$$

所以 $q_{i,K}$ 对 $K\subset J$ 来说有完全可加性.

**推论 1.2**　在定理 1.4 的条件下, 对任何 $i \neq j$, 恒有

(1) $\lim\limits_{t\to 0+} \dfrac{1}{t} p_{i,j}(t) = q_{i,j}$ 存在, 且 $0 \leqslant q_{i,j} < \infty$;

(2) 若记 $q_i$ 如定理 1.3 之极限, 则

$$\sum_{j\neq i} q_{i,j} \leqslant q_i.$$

**定理 1.5**　下列二条件等价:

(1) $\sup\limits_{i\in E} q_i < \infty$;

(2) $\lim\limits_{t\to 0+} \sup\limits_{i\in E}(1 - p_{i,i}(t)) = 0$.

若其中有一个条件满足, 则

$$q_i = \sum_{j\neq i} q_{i,j} \quad (\forall i \in E),$$

$$\lim_{t\to 0+} \frac{1 - p_{i,i}(t)}{t} = q_i \text{ 对 } i \in E \text{ 一致成立.}$$

**证**　(1) $\Rightarrow$ (2). 设 (1) 成立. 令 $c = \sup\limits_{i\in E} q_i < \infty$, 则由定理 1.3 有 $p_{i,i}(t) \geqslant \exp\{-q_i t\}$, 所以

$$p_{i,i}(t) \geqslant 1 - q_i t \geqslant 1 - ct,$$

从而 (2) 成立.

(2) $\Rightarrow$ (1). 设 (2) 成立, 则 $E \in \mathscr{G}$. 在定理 1.4 中取 $K = E - \{i\}$ 得

$$\lim_{t\to 0+} \frac{1 - p_{i,i}(t)}{t} = \lim_{t\to 0+} \frac{p_{i,K}(t)}{t} = q_{i,K}, \tag{1.23}$$

$0 \leqslant q_{i,K} < \infty$.

由定理 1.4 (2) 有

$$q_{i,K} = \sum_{j\neq i} q_{i,j},$$

但是, 由定理 1.3 有

$$\lim_{t\to 0+} \frac{1 - p_{i,i}(t)}{t} = q_i,$$

故

$$q_i = \sum_{j\neq i} q_{i,j} = q_{i,K} \quad (\forall i \in E). \tag{1.24}$$

由于 $E \in \mathscr{G}$, 所以由引理 1.2 得知对任意的 $0 < \varepsilon < \dfrac{1}{4}$, 必存在 $\tau = \tau(\varepsilon) > 0$ 使

$$1 - p_{j,j}(t) < \varepsilon \quad (\forall j \in E, 0 \leqslant t \leqslant \tau).$$

因此, 由引理 1.2 有

$$(1-4\varepsilon)\frac{p_{i,K}(u)}{u} \leqslant \frac{p_{i,K}(v)}{v} \quad (0<v\leqslant\tau). \tag{1.25}$$

在 (1.25) 中令 $u\to 0+$, 并注意 (1.23), 得

$$(1-4\varepsilon)q_{i,K} \leqslant \frac{p_{i,K}(v)}{v} \quad (0<v\leqslant\tau). \tag{1.26}$$

由 (1.24) 和 (1.26) 得知 $q_i = q_{i,K} \leqslant [(1-4\varepsilon)\tau]^{-1}$, 此即条件 (1) 成立.

剩下要证明的是: 在条件 (2) 成立时, 有

$$\lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t} = q_i, \quad \text{对 } i\in E \text{ 一致成立}.$$

由于

$$p_{i,i}(t) \geqslant e^{-q_i,t},$$

所以

$$q_i - \frac{1-p_{i,i}(t)}{t} \geqslant 0.$$

但是由 (1.26) 还有

$$\begin{aligned} q_i - \frac{1-p_{i,i}(t)}{t} = q_i - \frac{p_{i,K}(t)}{t} &\leqslant 4\varepsilon q_i \\ &\leqslant 4[(1-4\varepsilon)\tau]^{-1}\varepsilon \quad (\text{当 } 0<t\leqslant\tau \text{ 时}). \end{aligned}$$

所以

$$\lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t} = q_i, \quad \text{对 } i\in E \text{ 一致成立}.$$

定理 1.5 证毕.

**定理 1.6** 设 $P(t)=(p_{i,j}(t), i,j\in E)$ 是标准的转移矩阵, $q_i, q_{i,j}$ 的定义如定理 1.5. 令 $q_{i,i}=-q_i$. 固定任意一个 $i\in E$, 若 $0\leqslant q_i<\infty$, 则对任何 $j\in E, p_{i,j}(t)$ 在 $t\in(0,\infty)$ 内有连续微商 $p'_{i,j}(t)$, 而且它满足:

(1) $\sum\limits_{j\in E} p'_{i,j}(t) \leqslant 2q_i (\forall t\in[0,\infty))$;

(注意: $p'_{i,j}(t)$ 在 $(0,\infty)$ 连续, 但 $p'_{i,j}(0)=q_{i,j}$ 是没有问题的. (4) 说 $p'_{i,j}(t)$ 在 0 点右连续.)

(2) $\sum\limits_{j\in E} p'_{i,j}(t) = 0 (\forall t\in(0,\infty))$;

(3) $p'_{i,j}(s+t) = \sum\limits_{k\in E} p'_{i,k}(t)p_{k,j}(s) (\forall s\geqslant 0, t>0)$;

(4) $\lim\limits_{t\to 0+} p'_{i,j}(t) = p'_{i,j}(0) = q_{i,j}$;

(5) $\lim\limits_{t\to\infty} p'_{i,j}(t) = 0$.

由于 $q_i=0$ 时 $p_{i,j}(t)\equiv\delta_{i,j}$, 这时定理 1.6 显然成立. 下面设 $0<q_i<\infty$. 证明定理 1.6 以前, 先证明几条引理.

**引理 1.3** $e^{q_i t}p_{i,j}(t)$ 对 $t$ 来说单调非降.

**证** 任取 $s \geqslant 0, t \geqslant 0$, 由 $p_{i,i}(t) \geqslant e^{-q_i t}$ 和 (K − C) 方程式得

$$\begin{aligned} e^{q_i(s+t)}p_{i,j}(s+t) &\geqslant e^{q_i t}p_{i,i}(t)p_{i,j}(s)e^{q_i s} \\ &\geqslant e^{q_i s}p_{i,j}(s). \end{aligned}$$

**引理 1.4** 设 $f(x,y)$ 是二元实变实值函数, 若固定 $y, f(\bullet, y)$ 是 Lebesgue 可测的, 固定 $x, f(x, \bullet)$ 右连续, 则 $f(\bullet, \bullet)$ 是二元 Lebesgue 可测的.

证明甚易. 读者可作为习题验证之.

**引理 1.5** 在定理 1.6 的条件下, 存在唯一一组 $[0,\infty)$ 上的实值连续函数 $\{g_{i,j}(t), j \in E\}$, 使

$$p_{i,j}(t) = e^{-q_i t}\int_0^t q_i e^{q_i s}g_{i,j}(s)ds + \delta_{i,j}e^{-q_i t}, \tag{1.27}$$

而且还满足:

(1) $g_{i,j}(t) \geqslant 0 (\forall t \geqslant 0), \sum\limits_{j\in E} g_{i,j}(t) \equiv 1 (\forall t > 0)$;

(2) $g_{i,j}(s+t) = \sum\limits_{k\in E} g_{i,k}(s)p_{k,j}(t) (\forall s > 0, t \geqslant 0)$;

(3) $\lim\limits_{t\to 0+} g_{i,j}(t) = \begin{cases} 0, & \text{当 } j = i, \\ q_{i,j}/q_i, & \text{当 } j \neq i; \end{cases}$

(4) $\lim\limits_{t\to\infty} g_{i,j}(t) = \lim\limits_{t\to\infty} p_{i,j}(t) = \pi_{i,j}$ 存在.

**证** 分几步来证明此引理.

第一步. 存在一组 $[0,\infty)$ 上的 Lebesgue 可测函数 $\{r_{i,j}(t), j \in E\}$ 使

$$r_{i,j}(t) \geqslant 0, \quad \sum_{j\in E} r_{i,j}(t) \equiv 1 \quad (\forall t \geqslant 0, j \in E), \tag{1.28}$$

而且

$$e^{q_i t}p_{i,j}(t) = \int_0^t q_i e^{q_i s} r_{i,j}(s)ds + \delta_{i,j} \quad (t \geqslant 0, j \in E). \tag{1.29}$$

事实上, 由定理 1.1 及引理 1.3 得知: 对任何 $J \subset E, e^{q_i, t}p_{i,J}(t)$ 对 $t$ 来说单调非降且绝对连续, 所以存在非负的 Lebesgue 可测函数 $r_{i,j}(s)$ 和 $h_{i,j}(t)$, 使

$$e^{q_i t}p_{i,j}(t) = \int_0^t q_i e^{q_i s} r_{i,j}(s)ds + \delta_{i,j}, \tag{1.30}$$

$$e^{q_i t}p_{i,E-\{j\}}(t) = \int_0^t q_i e^{q_i s} h_{i,j}(s)ds + p_{i,E-\{j\}}(0), \tag{1.31}$$

把上面两式相加得

$$e^{q_i t} = \int_0^t q_i e^{q_i s}(r_{i,j}(s) + h_{i,j}(s))ds + 1,$$

所以

$$r_{i,j}(s) + h_{i,j}(s) = 1, \quad \text{a.s.}.$$

但是 $h_{i,j}(s) \geqslant 0$, 所以

$$0 \leqslant r_{i,j}(s) \leqslant 1, \quad \text{a.s.}.$$

式 (1.30) 左右两边对 $j \in E$ 分别求和得

$$e^{q_i t} = \int_0^t q_i e^{q_i s} \sum_{j \in E} r_{i,j}(s)ds + 1,$$

所以

$$\sum_{j \in E} r_{i,j}(s) = 1, \quad \text{a.s.}.$$

不妨认为

$$\sum_{j \in E} r_{i,j}(s) \equiv 1, \quad r_{i,j}(s) \geqslant 0. \tag{1.32}$$

第二步. 定义 $g_{i,j}(t)$.

以 (1.30) 代入

$p_{i,j}(s+t) = \sum_{k \in E} p_{i,k}(s)p_{k,j}(t) \quad (s \geqslant 0, t \geqslant 0)$, 得

$$\begin{aligned}
& q_i e^{-q_i(s+t)} \int_0^{s+t} e^{q_i u} r_{i,j}(u)du + e^{-q_i(s+t)}\delta_{i,j} \\
&= p_{i,j}(s+t) = \sum_{k \in E} p_{i,k}(s)p_{k,j}(t) \\
&= \sum_{k \in E} q_i e^{-q_i s} \int_0^s e^{q_i u} r_{i,k}(u)p_{k,j}(t)du + e^{-q_i s}p_{i,j}(t).
\end{aligned} \tag{1.33}$$

由 (1.33) 和 (1.30) 可得

$$\begin{aligned}
& e^{-q_i s} \int_0^s e^{q_i u} \sum_{k \in E} r_{i,k}(u)p_{k,j}(t)du \\
&= \frac{1}{q_i}(p_{i,j}(s+t) - e^{-q_i s}p_{i,j}(t)) \\
&= e^{-q_i(s+t)}\left[\int_0^{s+t} e^{q_i u} r_{i,j}(u)du - \int_0^t e^{q_i u} r_{i,j}(u)du\right] \\
&= e^{-q_i(s+t)} \int_t^{s+t} e^{q_i u} r_{i,j}(u)du \\
&= e^{-q_i s} \int_0^s e^{q_i u} r_{i,j}(u+t)du.
\end{aligned} \tag{1.34}$$

因此, 对每个 $t \geqslant 0$, 存在一个 Lebesgue 零测集 $A_t$, 使

$$r_{i,j}(s+t)=\sum_{k\in E} r_{i,k}(s)p_{k,j}(t) \quad (\text{当 } s\overline{\in}A_t, t\geqslant 0). \tag{1.35}$$

由于 (1.35) 右边级数每一项都是 $s$ 的 Lebesgue 可测函数 (固定 $t$) 又是 $t$ 的连续函数 (固定 $s$), 所以由引理 1.4 得知 $r_{i,j}(s+t)$ 是 $(s,t)$ 的二元 Lebesgue 可测函数, 从而

$$B=\left\{(s,t): s>0, t>0, r_{i,j}(s+t)\neq \sum_{k\in E} r_{i,k}(s)p_{k,j}(t)\right\}$$

是二维 Lebesgue 可测集. 再根据 Fubini 定理和式 (1.35) 可知

$$L_2(B)=\int_0^{\infty} L_1(A_t^c)dt=0, \tag{1.36}$$

其中 $L_k$ 表示 $k$ 维 Lebesgue 测度, $A_t^c$ 是 $A_t$ 的补集.

作线性变换 $u=s, v=s+t$, 它把集合 $B$ 变为 $D$. 由于线性变换是保测的, 所以

$$L_2(D)=L_2(B)=0.$$

由 $B$ 的定义可看出

$$D=\left\{(u,v): v>u>0, r_{i,j}(v)\neq \sum_{k\in E} r_{i,k}(u)p_{k,j}(v-u)\right\}.$$

记

$$D_u=\{v: v>0, (u,v)\in D\}$$

为 $D$ 在 $u$ 的截口集. 由 Fubini 定理及 $L_2(D)=0$ 得知: 存在一个 $u$ 的集合 $H$, 使 $L_1(H)=0$ 且 $u\overline{\in}H$ 时 $L_1(D_u)=0$.

定义 $g_{i,j}(v)$ 如下:

当 $v>0$ 时, 取 $w<v, w\overline{\in}H$, 令

$$g_{i,j}(v)=\sum_{k\in E} r_{i,k}(w)p_{k,j}(v-w), \tag{1.37}$$

当 $v=0$ 时, 令

$$g_{i,j}(0)=\lim_{v\to 0+} g_{i,j}(v). \tag{1.38}$$

第三步. 证明上面定义的 $g_{i,j}$ 即为所求.

首先说明上面定义的 $g_{i,j}$ 不依赖 $w$ 的选取. 事实上, 若取 $w_i < v, v > 0$, $w_i \overline{\in} H$, 则由 $H$ 的选取可知

$$\sum_{k\in E} r_{i,k}(w_1)p_{k,j}(t-w_1) = r_{i,j}(t)$$

$$= \sum_{k\in E} r_{i,k}(w_2)p_{k,j}(t-w_2) \tag{1.39}$$

对 $(\max(w_1,w_2),\infty)$ 中 a.s. 的 $t$ 成立. 仿定理 1.2 可证 (1.39) 左右两边都是 $t$ 的连续函数, 所以 (1.39) 对一切 $t \in (\max(w_1,w_2),\infty)$ 都成立, 特别地

$$\sum_{k\in E} r_{i,k}(w_1)p_{k,j}(v-w_1) = \sum_{k\in E} r_{i,k}(w_2)p_{k,j}(v-w_2),$$

此即 $g_{i,j}$ 的定义不依赖于 $w$ 的选取.

取 $a>0, v>a$, 由 $p_{k,j}(\bullet)$ 连续, $r_{i,k}(\bullet)$ 满足 (1.32) 及

$$g_{i,j}(v) = \sum_{k\in E} r_{i,k}(a)p_{k,j}(v-a),$$

仿定理 1.2 可证 $g_{i,j}(v)$ 在 $v>a$ 连续. 由 $a>0$ 可任意小知 $g_{i,j}(v)$ 在 $v>0$ 连续. 由定义知 $g_{i,j}(v)$ 在 0 处右连续.

由 $g_{i,j}(v) = r_{i,j}(v)$, a.s. 及 $r_{i,j}$ 满足 (1.27) 知 $g_{i,j}(v)$ 满足 (1.27).

最后证明 $g_{i,j}$ 满足引理 1.5 中的条件 (1)~(4).

(1) 可由 (1.28)、(1.40) 及 $g_{i,j}$ 的连续性可得.

(2) 由 (1.35) 和 $g_{i,j}(v)$ 的连续性及

$$g_{i,j}(v) = r_{i,j}(v), \quad \text{a.s.},$$

易证条件 (2) 成立.

(3) 由 (2) 知:

$$g_{i,j}(s+t) \geqslant g_{i,j}(s)p_{j,j}(t) \quad (s>0, t>0).$$

令 $s \to 0+$ 得

$$g_{i,j}(t) \geqslant \left(\limsup_{s\to 0+} g_{i,j}(s)\right) p_{j,j}(t),$$

从而由 $P(t)$ 的标准性知:

$$\liminf_{t\to 0+} g_{i,j}(t) \geqslant \limsup_{s\to 0+} g_{i,j}(s),$$

所以 $\lim\limits_{t\to 0+} g_{i,j}(t)$ 存在.

由 (1.27) 得

$$
\begin{aligned}
q_i &= \lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t} \\
&= q_i - q_i\lim_{t\to 0+} g_{i,i}(t).
\end{aligned}
\tag{1.40}
$$

但是 $q_i>0$, 所以

$$
\lim_{t\to 0+} g_{i,i}(t)=0. \tag{1.41}
$$

而当 $i\neq j$ 时, 由 (1.27) 有

$$
q_{i,j}=\lim_{t\to 0+}\frac{p_{i,j}(t)}{t}=q_i\lim_{t\to 0+} g_{i,j}(t),
$$

所以

$$
\lim_{t\to 0+} g_{i,j}(t)=\frac{q_{i,j}}{q_i}. \tag{1.42}
$$

综合 (1.41)、(1.42) 知条件 (3) 成立.

(4) 由 (1)、(2) 及 $\lim\limits_{t\to\infty} p_{k,j}(t)=\pi_{k,j}$ 存在并用控制收敛定理可证:

$$
\lim_{t\to\infty} g_{i,j}(t)\ \text{存在}.
$$

在 (1.27) 式中令 $t\to\infty$, 并注意 $\lim\limits_{t\to\infty} p_{i,j}(t)=\pi_{i,j}$ 可知 $\lim\limits_{t\to\infty} g_{i,j}(t)=\pi_{i,j}$.

$\{g_{i,j}, j\in E\}$ 的唯一性是显然的. 引理证毕.

现在我们用上述诸引理来证明定理 1.6. 由引理 1.5 的 (1.27) 式知 $p_{i,j}(t)$ 在 $(0,\infty)$ 内有连续微商 $p'_{i,j}(t)$. 下面逐一验证 $p'_{i,j}(t)$ 满足条件 (1)~(5).

对 (1.27) 求微商得

$$
q_i e^{q_i t} p_{i,j}(t) + e^{q_i t} p'_{i,j}(t) = q_i e^{q_i t} g_{i,j}(t) \quad (t>0). \tag{1.43}
$$

由 (1.43) 得

(1)
$$
\begin{aligned}
\sum_{j\in E}|p'_{i,j}(t)| &\leqslant q_i\sum_{k\in E}(g_{i,j}(t)+p_{i,j}(t)) \\
&\leqslant 2q_i \quad (\forall t>0);
\end{aligned}
$$

(2)
$$
\begin{aligned}
\sum_{j\in E} p'_{i,j}(t) &= q_i\left(\sum_{j\in E} g_{i,j}(t)-\sum_{j\in E} p_{i,j}(t)\right) \\
&= 0 \quad (\forall t>0);
\end{aligned}
$$

(3) 当 $s\geqslant 0, t>0$ 时, 有

$$
g_{i,j}(s+t)=\sum_{k\in E}^{\infty} g_{i,k}(s)p_{k,j}(t),
$$

$$
p_{i,j}(s+t)=\sum_{k\in E}^{\infty} p_{i,k}(s)p_{k,j}(t),
$$

把上述两式相减并应用 (1.27) 式得

$$\begin{aligned} p'_{i,j}(s+t) &= q_i(g_{i,j}(s+t)-p_{i,j}(s+t)) \\ &= q_i\sum_{k\in E}(g_{i,k}(s)-p_{i,k}(s))p_{k,j}(t) \\ &= \sum_{k\in E}p'_{i,k}(s)p_{k,j}(t). \end{aligned}$$

(4) 可由 (1.27) 直接得到.

(5) 由 (1.27) 还有

$$\begin{aligned} \lim_{t\to\infty}p'_{i,j}(t) &= \lim_{t\to\infty}q_i(g_{i,j}(t)-p_{i,j}(t)) \\ &= q_i(\pi_{i,j}-\pi_{i,j})=0. \end{aligned}$$

定理 1.6 证毕.

**定理 1.7** 设 $P(t)=(p_{i,j}(t),i,j\in E)$ 是标准的转移矩阵, $q_i,q_{i,j}$ 如前所定义. 假设 $q_i<\infty$, 则下列二条件等价:

(1) $p'_{i,j}(t)=\sum\limits_{k\in E}q_{i,k}p_{k,j}(t) \quad (t\geqslant 0,j\in E)$;

(2) $\sum\limits_{k\in E}q_{i,k}=0$.

**证** (1) ⇒ (2). 设 (1) 成立. 由定理 1.6 知: 对任何 $t>0$, 有

$$\begin{aligned} 0=\sum_{j\in E}p'_{i,j}(t) &= \sum_{j\in E}\sum_{k\in E}q_{i,k}p_{k,j}(t) \\ &= \sum_{k\in E}q_{i,k}\sum_{j\in E}p_{k,j}(t)=\sum_{k\in E}q_{i,k}. \end{aligned}$$

(2)⇒(1). 设 (2) 成立. 当 $q_i=0$ 时, (1) 显然成立. 下设 $0<q_i<\infty$. 由引理 1.5 有:

$$g_{i,j}(s+t)=\sum_{k\in E}g_{i,k}(s)p_{k,j}(t) \quad (s>0,t\geqslant 0),$$

且

$$\lim_{s\to 0+}g_{i,k}(s)=\begin{cases}0, & 当\ k=i,\\ q_{i,k}/q_i, & 当\ k\neq i,\end{cases}$$

又因为 $g_{i,j}(s)$ 在 $[0,\infty)$ 上连续, $\sum\limits_{j\in E}g_{i,j}(s)=1(\forall s>0)$. 故由 (2) 可得

$$\sum_{k\in E}\lim_{s\to 0+}g_{i,k}(s)=1.$$

所以由 (1.27) 式及第一章定理 5.8 (Helly 定理)

$$\begin{aligned}p'_{i,j}(t) &= q_i(g_{i,j}(t) - p_{i,j}(t)) \\ &= q_i\left(\lim_{s\to 0+}\sum_{k\in E} g_{i,k}(s)p_{k,j}(t) - p_{i,j}(t)\right) \\ &= q_i\left(\sum_{k\in E}\left(\lim_{s\to 0+} g_{i,k}(s)\right)p_{k,j}(t) - p_{i,j}(t)\right) \\ &= q_i\left(\sum_{k\neq i}\frac{q_{i,k}}{q_i}p_{k,j} - p_{i,j}(t)\right) \\ &= \sum_{k\in E} q_{i,k}p_{k,j}(t) \quad (\forall t > 0).\end{aligned}$$

而当 $t=0$ 时上式显然成立. 定理证毕.

**定理 1.8**　设 $P(t) = (p_{i,j}(t), i,j\in E)$ 是标准的转移矩阵, $Q = (q_{i,j}, i,j \in E), q_i, q_{i,j}$ 之定义如前. 若

$$\sup_{i\in E} q_i \leqslant c < \infty,$$

则

(1) $P'(t) = QP(t) = P(t)Q \quad (t \geqslant 0)$;

(2) $P(t) = e^{Qt} \stackrel{\text{def.}}{=\!=} \sum\limits_{k=0}^{\infty}\dfrac{(Qt)^k}{k!} \quad (t\geqslant 0)$;

(3) $Q\Pi = \Pi Q = 0$, 其中

$$\Pi = (\pi_{i,j}, i,j\in E), \quad \pi_{i,j} = \lim_{t\to\infty} p_{i,j}(t).$$

反之任给矩阵 $Q = (q_{i,j}, i,j\in E)$, 只要 $-\infty < q_{i,i} \leqslant 0, 0 \leqslant q_{i,j} < \infty (i \neq j, i,j \in E)$, $\sum\limits_{j\in E} q_{i,j} = 0 (\forall i\in E)$, $\sup\limits_{i\in E}(-q_{i,i}) \leqslant c < \infty$, 则 $P(t) \stackrel{\text{def.}}{=\!=} e^{Qt}$ 是一个标准的转移矩阵.

**证**　(1) 因为 $\sup\limits_{i\in E} q_i \leqslant c < \infty$, 所以由定理 1.5 知

$$\sum_{j\in E} q_{i,j} = 0 \quad (\forall i \in E).$$

因此, 由定理 1.7 得

$$P'(t) = QP(t) \quad (\forall t \geqslant 0).$$

而对任何 $t>0, 0<h<t, i,j\in E$, 总有

$$\frac{p_{i,j}(t+h)-p_{i,j}(t)}{h}=\frac{\sum\limits_{k\in E}p_{i,k}(t)(p_{k,j}(h)-\delta_{i,j})}{h}, \tag{1.44}$$

$$\frac{p_{i,j}(t)-p_{i,j}(t-h)}{h}=\frac{\sum\limits_{k\in E}p_{i,k}(t-h)(p_{k,j}(h)-\delta_{i,j})}{h}, \tag{1.45}$$

但是由定理 1.5 及 $\sup\limits_{i\in E}q_i\leqslant c<\infty$ 知:

$$\lim_{h\to 0+}\frac{1-p_{i,i}(h)}{h}=q_i,\quad 对\ i\in E\ 一致成立,$$

所以对任何固定的 $\varepsilon>0$, 存在 $h_0$ 与 $k$ 无关使

$$\frac{1-p_{k,k}(h)}{h}<q_k+\varepsilon\leqslant c+\varepsilon\quad (当\ h\leqslant h_0, k\in E), \tag{1.46}$$

更有

$$\frac{p_{k,j}(h)}{h}\leqslant\frac{1-p_{k,k}(h)}{h}\leqslant c+\varepsilon\quad (当\ h\leqslant h_0, k\neq j). \tag{1.47}$$

在 (1.44) 和 (1.45) 中令 $h\to 0+$, 并利用控制收敛定理可得:

$$p'_{i,j}(t)=\sum_{k\in E}p_{i,k}(t)q_{k,j}\quad (t>0, i,j\in E).$$

显然 $t=0$ 时上式亦成立. 总之 $P'(t)=P(t)Q$ 得证. 仿之可证 $P'(t)=QP(t)$. (1) 证毕.

(2) 由 (1) 有

$$p'_{i,j}(t)=\sum_{k\in E}q_{i,k}p_{k,j}(t)\quad (i,j\in E, t\geqslant 0), \tag{1.48}$$

由定理 1.6 及 $\sup\limits_{i\in E}q_i\leqslant c<\infty$ 有

$$\sum_{k\in E}\left|q_{i,k}p'_{k,j}(t)\right|\leqslant\sum_{k\in E}|q_{i,k}|2q_k\leqslant 4c^2. \tag{1.49}$$

所以 $\sum\limits_{k\in E}q_{i,k}p'_{k,j}(t)$ 在 $[0,\infty)$ 上一致收敛, 从而

$$p'_{i,j}(t)=\sum_{k\in E}q_{i,k}p_{k,t}(t)$$

在 $[0,\infty)$ 上可以逐项微商. 微商之并用 (1) 得

$$\begin{aligned}p''_{i,j}(t)&=\sum_{k\in E}q_{i,k}p'_{k,j}(t)\\&=\sum_{k\in E}q_{i,k}\sum_{l\in E}q_{k,l}p_{l,j}(t)\quad (t\geqslant 0, i,j\in E).\end{aligned}$$

此即

$$P''(t) = Q^2 P(t) \quad (t \geqslant 0).$$

仿之, 继续作下去可得

$$\frac{d^n P(t)}{dt^n} = Q^n P(t) \quad (n \geqslant 0, t \geqslant 0). \tag{1.50}$$

所以

$$\begin{aligned} P(t) &= \sum_{n=0}^{\infty} \left( \frac{d^n P(t)}{dt^n} \right) \Bigg|_{t=0} \frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!} = e^{Qt}. \end{aligned}$$

(3) 因为

$$P'(t) = QP(t), \quad \lim_{t\to\infty} P(t) = \Pi, \quad \lim_{t\to\infty} P'(t) = 0,$$

所以用控制收敛定理可得:

$$\begin{aligned} 0 &= \lim_{t\to\infty} P'(t) = \lim_{t\to\infty} QP(t) = Q\left(\lim_{t\to\infty} P(t)\right) \\ &= Q\Pi. \end{aligned}$$

又因为

$$P'(t) = P(t)Q,$$

若记

$$D_q = \mathrm{diag}(q_i, i \in E)$$

为主对角线上的元素为 $\{q_i, i \in E\}$ 的对角矩阵, $S \stackrel{\text{def.}}{=\!=} Q + D_q$ 是非负矩阵, 则根据 Fatou 引理可得

$$0 = \liminf_{t\to\infty} P'(t) \geqslant \Pi S - \Pi D_q.$$

但是 $\sup\limits_{i\in E} q_i \leqslant c < \infty$, 所以 $Q\mathbf{1} = 0$ (其中 $\mathbf{1}$ 是一切元素皆为 1 的列向量, 下同.) 因此

$$\begin{aligned} (\Pi S - \Pi D_q)\mathbf{1} &= \Pi S\mathbf{1} - \Pi D_q\mathbf{1} \\ &= \Pi D_q\mathbf{1} - \Pi D_q\mathbf{1} = 0. \end{aligned}$$

故 $\Pi Q = 0$.

现在任给一个满足定理 1.8 中全部条件的矩阵 $Q = (q_{i,j}, i, j \in E)$. 要证 $P(t) \stackrel{\text{def.}}{=\!=} e^{Qt}$ 是一个标准的转移矩阵. 显然 $P(t)\mathbf{1} = \mathbf{1}, P(s+t) = P(s)P(t)(\forall s, t \geqslant$

0), 而且 $\lim\limits_{t\to 0+} P(t) = I$ ($I$ 是单位矩阵). 下面只需证明: $P(t) \geqslant 0$ (对一切 $t \in [0,\infty)$), 则定理获证. 先设

$$q_{i,j} > 0 \quad (\text{对一切 } i \neq j, i, j \in E).$$

事实上, 若令 $Q^n = (q_{i,j}^{(n)}, i, j \in E)$, 则由 $\sup\limits_{i\in E}(-q_{i,i}) \leqslant c < \infty$ 可得

$$\left|q_{i,j}^{(2)}\right| \leqslant \sum_{k\in E} |q_{i,k}|\,|q_{k,j}| \leqslant c \sum_{k\in E} |q_{i,k}| \leqslant 2c^2,$$

仿之可证: 对任何 $n \geqslant 1$ 有

$$\left|q_{i,j}^{(n)}\right| \leqslant 2^{n-1}c^n \quad (\forall i, j \in E). \tag{1.51}$$

由于

$$P(t) = e^{Qt} \doteq I + Qt + \frac{1}{2!}(Qt)^2 + \cdots, \tag{1.52}$$

所以若取 $\delta > 0$ 充分小使

$$2c\delta + \frac{(2c\delta)^2}{2!} + \frac{(2c\delta)^3}{3!} + \cdots = e^{2c\delta} - 1 < \frac{1}{2}, \tag{1.53}$$

则由 (1.51)、(1.52)、(1.53) 有

$$\begin{aligned} p_{i,i}(t) &= 1 + q_{i,i}t + \frac{1}{2!}q_{i,i}^{(2)}t^2 + \cdots \\ &\geqslant 1 - \left(\sum_{k=1}^{\infty} \frac{1}{k!}\left|q_{i,i}^{(k)}\right| \delta^k\right) \\ &\geqslant \frac{1}{2} \quad (\forall i \in E, t \in [0,\delta]). \end{aligned} \tag{1.54}$$

任取 $i \neq j$, 由 $P'(0) = Q$ 及

$$p'_{i,j}(0) = q_{i,j} > 0, \quad p_{i,j}(0) = 0$$

得知: 存在 $\varepsilon_{i,j} > 0$, 使

$$p_{i,j}(t) \geqslant 0 \quad (\forall t \in [0, \varepsilon_{i,j}]).$$

取 $\eta_{i,j} = \min(\delta, \varepsilon_{i,j})$, 则

$$p_{i,j}(2t) \geqslant p_{i,i}(t)p_{i,j}(t) \geqslant 0 \quad (i \neq j, t \in [0, \eta_{i,j}]).$$

所以

$$p_{i,j}(t) \geqslant 0 \quad (i \neq j, t \in [0,\infty)). \tag{1.55}$$

由 (1.54) 及 $P(2t) = P(t)P(t)$ 得

$$p_{i,i}(t) \geqslant 0 \quad (\forall i \in E, t \in [0, \infty)). \tag{1.56}$$

由 (1.55)、(1.56) 可知:

$$\text{当 } q_{i,j} > 0 \ \ (i \neq j) \ \ \text{时} \ \ P(t) \geqslant 0 (\forall t \in [0, \infty)).$$

现在取消这一假设. 构造

$$\begin{aligned} Q_n &= (r_{i,j}^{(n)}, i, j \in E), \\ r_{i,j}^{(n)} &= q_{i,j} + \frac{a_{i,j}}{n}(i \neq j), \\ r_{i,i}^{(n)} &= q_{i,i} - \frac{a_i}{n}(i \in E), \end{aligned}$$

其中 $a_{i,j} > 0, \sum\limits_{j \neq i} a_{i,j} = a_i, \sup\limits_{i \in E} a_i \leqslant d < 1$, 则 $Q_n$ 满足定理 1.8 中 $Q$ 的全部条件, 且 $r_{i,j}^{(n)} > 0$ (对一切 $i \neq j$), 所以

$$P_n(t) \stackrel{\text{def.}}{=\!=} e^{Q_n t} \geqslant 0 \quad (\forall t \in [0, \infty)),$$

故 $P(t) = \lim\limits_{n \to \infty} P_n(t) \geqslant 0$. 定理 1.8 证毕.

**思考题** 前面得到的八个定理, 是关于标准转移矩阵的. 对标准的准转移矩阵, 是否有类似的结果? 如果有, 它们的相应的结果是什么?

提示: 若 $P(t) = (p_{i,j}(t), i, j \in E)$ 是准转移矩阵, 取 $\Delta \overline{\in} E$, 考虑扩展矩阵

$$P_\Delta(t) = \begin{array}{c} \\ E \\ \Delta \end{array} \!\!\begin{array}{c} \begin{array}{cc} E & \Delta \end{array} \\ \begin{pmatrix} P(t) & d(t) \\ 0 & 1 \end{pmatrix} \end{array}, \quad \text{其中 } d(t) = \mathbf{1} - P(t)\mathbf{1} \text{ 是列向量.}$$

## §2 $Q$ 过程的存在唯一性

本节沿袭 §1 的符号. 在 §1 中我们系统地研究了标准的转移矩阵 $P(t)$ 的连续性及可微性. 在最后的思考题中, 曾问到对标准的准转移矩阵是否有类似的结果? 其实在提示中给出的扩展矩阵 $P_\Delta(t)$ 是一个标准的转移矩阵, $P(t)$ 是它在 $E \times E$ 上的局限. 若 $P_\Delta(t)$ 有连续性、可微性, $P(t)$ 自然也有, 只是某些方程式的形式略有改变.

这一节我们主要研究可微性的逆问题, 即是给定某种矩阵 $Q = (q_{i,j}, i, j \in E)$, 是否恒存在一个标准的准转移矩阵 $P(t)$, 使 $P'(0) = Q$? 如果有, 是否唯一? 如果不一定唯一, 那么唯一的充分必要条件是什么?

**定义 2.1** 称矩阵 $Q=(q_{i,j},i,j\in E)$ 是一个**转移强度矩阵**, 如果

(1) $0\geqslant q_{i,i}\geqslant -\infty(\forall i\in E)$;

(2) $0\leqslant q_{i,j}<\infty(\forall i\neq j,i,j\in E)$;

(3) $\sum\limits_{j\neq i}q_{i,j}\leqslant -q_{i,i}$.

特别地, 若转移强度矩阵 $Q$ 满足

$$0\geqslant q_{i,i}>-\infty \quad (\forall i\in E), \tag{2.1}$$

则称 $Q$ 是**全稳定的**转移强度矩阵. 更特别地, 若全稳定的转移强度矩阵 $Q$ 还满足

$$\sum_{j\in E}q_{i,j}=0 \quad (\forall i\in E), \tag{2.2}$$

则称 $Q$ 是**保守的**转移强度矩阵.

本书仅研究全稳定的转移强度矩阵 $Q$, 所以 "全稳定" 三字略去不写.

只要存在一个 $i_0\in E$, 使

$$q_{i_0,i_0}=-\infty \tag{2.3}$$

(这时称 $i_0$ 是 $Q$ 的一个瞬时状态), 则 §1 中的可微性的逆问题, 所谓 $Q$ 过程问题就变得复杂得多. 若 $0\geqslant q_{i_0,i_0}>-\infty$, 则称 $i_0$ 是 $Q$ 的一个稳定状态; 称 $i_0$ 是**吸收状态**, 若 $q_{i_0,i_0}=0$.

为什么满足 $q_{i_0,i_0}=0$ 的 $i_0$ 称为吸收状态, 满足 $q_{i_0,i_0}=-\infty$ 的 $i_0$ 称为**瞬时状态**, 而满足 $0\geqslant q_{i_0,i_0}>-\infty$ 的 $i_0$ 称为**稳定状态**? 这要用 Markov 过程的轨道性质才能解释.

对于时齐的可分的以 $P(t)=(p_{i,j}(t),i,j\in E)$ 为转移矩阵的 Markov 过程 $\{X_t,t\in[0,\infty)\}$ 来说, 总有

$$\begin{aligned}&P\left(X_u=i,\ \text{对一切}\ s\leqslant u\leqslant s+t|X_s=i\right)\\&=e^{-q_it}(s\geqslant 0,t>0,i\in E,e^{-\infty}\overset{\text{def.}}{=\!=}0).\end{aligned} \tag{2.4}$$

其中 $q_i$ 如定理 1.3 所定义.

(2.4) 的证明可参见 [13] P.152 定理 5.

(1) 对任何标准的转移矩阵 $P(t)=(p_{i,j}(t),i,j\in E)$ 来说, $P'(0)=Q$ 是一个转移强度矩阵.

(2) 由 (2.4) 看出: 当 $i$ 是瞬时状态, 即 $q_i=\infty$, 对任何 $s\geqslant 0,t>0,P(X_u=i,\forall s\leqslant u\leqslant s+t|X_s=i)=0$, 即是从状态 $i$ 出发, 在 $i$ 逗留一段正时间的概率为 0.

(3) 由 (2.4) 看出: 当 $i$ 是吸收状态时, 即 $q_i = 0$, 则由状态 $i$ 出发, 过程永远停留在 $i$.

(4) 当 $0 < i < \infty$, 即 $i$ 是非吸收的稳定状态时, 由 (2.4) 看出: 由状态 $i$ 出发, 逗留在 $i$ 的时间服从负指数分布.

由 (1)~(4) 看出: 上述诸定义是符合概率直观的.

**定义 2.2** 给定一个转移强度矩阵 $Q = (q_{i,j}, i, j \in E)$, 称标准的准转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$ 是一个 $Q$ 过程, 如果

$$P'(0) = Q.$$

特别地, 满足 $P(t)\mathbf{1} = \mathbf{1}$ 的 $Q$ 过程称为**不间断的**, 反之称为**间断的**. 若 $P'(t) = QP(t)$, 则称 $P(t)$ 满足**倒退方程** (简称为满足 (B)); 若 $P'(t) = P(t)Q$, 则称 $P(t)$ 满足**前进方程** (简称为满足 (F)).

**定义 2.3** 设 $P(t)$ 是任意一个标准的准转移矩阵, 称

$$R(\lambda) = \int_0^\infty e^{-\lambda t} P(t) dt \quad (\lambda > 0) \tag{2.5}$$

为 $P(t)$ 的 **Laplace 变换**.

注意: 由 $P(t) \geqslant 0, P(t)\mathbf{1} \leqslant \mathbf{1}, P(t)$ 关于 $t$ 有连续性, 所以 (2.5) 右方之积分存在且有穷 (对每个 $\lambda > 0$).

**定理 2.1** 设 $P(t)$ 是标准的准转移矩阵,$R(\lambda)$ 是其 Laplace 变换, 则 $R(\lambda)$ 满足:

(1) 正则化条件:

$$R(\lambda) \geqslant 0, \lambda R(\lambda)\mathbf{1} \leqslant \mathbf{1};$$

(2) 预解方程式:

$$R(\lambda) - R(\mu) + (\lambda - \mu)R(\lambda)R(\mu) = 0 \quad (\lambda, \mu > 0);$$

(3) 连续性条件:

$$\lim_{\lambda \to \infty} \lambda R(\lambda) = I.$$

特别地, 若 $R(\lambda)$ 是 $Q$ 过程 $P(t)$ 的 Laplace 变换, 则除了 (1) 和 (2) 仍然成立以外, (3) 还可以加强为:

(3)′
$$\lim_{\lambda \to \infty} \lambda(\lambda R(\lambda) - I) = Q.$$

显然 (3)′ 蕴涵了 (3).

证明甚易, 读者可作为习题验证之.

**附注 2.1** 满足 (1)、(2)、(3) 的 $R(\lambda)$ 必是某个标准的准转移矩阵的 Laplace 变换; 满足 (1)、(2)、(3)′ 的 $R(\lambda)$ 必是某个 $Q$ 过程的 Laplace 变换.

附注2.1 的证明较难, 有兴趣的读者可参见 [37] 第二编定理 2.1 和定理 2.2. 那里是对一般状态空间来证明的, 即 $(E,\mathscr{E})$ 是任一抽象的可测空间.

**定理 2.2** 设 $Q=\{q_{i,j},i,j\in E\}$ 是一个转移强度矩阵, $P(t)$ 是一个 $Q$ 过程, $q_i=-q_{i,i},D_q=\mathrm{diag}(q_i,i\in E),S=Q+D_q,\varDelta(t)=\mathrm{diag}(e^{-q_it},i\in E)$, 则

(1) $P'(t)\geqslant QP(t)\quad(\forall t\geqslant 0)$;

(2) $P'(t)\geqslant P(t)Q\quad(\forall t\geqslant 0)$,

和

(1)′ $P(t)\geqslant \varDelta(t)+\int_0^t\varDelta(t-s)SP(s)ds$;

(2)′ $P(t)\geqslant \varDelta(t)+\int_0^t\varDelta(t-s)(D_qP(s)+P(s)S-P(s)D_q)ds$.

**证** (1) 和 (2) 用 $P'(t)$ 的定义及 Fatou 引理立即可得.

(1)′ 和 (2)′ 用分部积分法并用 (1)、(2) 立即可算出.

**定理 2.3** 沿用定理 2.2 的符号. 下列三条件等价:

(B): $P'(t)=QP(t)\quad(\forall t\geqslant 0)$;

(B)′: $P(t)=\varDelta(t)+\int_0^t\varDelta(t-s)SP(s)ds\quad(\forall t\geqslant 0)$;

($\mathrm{B}_\lambda$): $(\lambda I-Q)R(\lambda)=I\quad(\forall\lambda>0)$,

其中 $R(\lambda)$ 是 $P(t)$ 的 Laplace 变换. 类似地, 下列三条件亦等价:

(F): $P'(t)=P(t)Q\quad(\forall t\geqslant 0)$;

(F′): $P(t)=\varDelta(t)+\int_0^t\varDelta(t-s)(D_qP(s)+P(s)S-P(s)D_q)ds\quad(\forall t\geqslant 0)$;

($\mathrm{F}_\lambda$): $R(\lambda)(\lambda I-Q)=I\quad(\forall\lambda>0)$.

**证** 由分部积分法可知

$$P(t)=\varDelta(t)+\int_0^t\varDelta(t-s)(P'(s)+D_qP(s))ds. \tag{2.6}$$

由 (2.6) 立即看出: (B) ⇔ (B′); (F) ⇔ (F′).

下面证明: (F′) ⇔ ($\mathrm{F}_\lambda$)((B′) ⇔ ($\mathrm{B}_\lambda$) 类似).

容易算出 $\Delta(t)$ 的 Laplace 变换为 $(\lambda I + D_q)^{-1}$. 又因为

$$\begin{aligned}
&\int_0^\infty e^{-\lambda t}dt\int_0^t \Delta(t-s)(D_qP(s)+P(s)S-P(s)D_q)ds\\
=&\int_0^\infty ds\left(\int_0^\infty e^{-\lambda t}e^{-\lambda s}\Delta(t)dt(P(s)S+D_qP(s)-P(s)D_q)\right)\\
=&\int_0^\infty (\lambda I+D_q)^{-1}e^{-\lambda s}(P(s)S+D_qP(s)-P(s)D_q)ds\\
=&(\lambda I+D_q)^{-1}(R(\lambda)S+D_qR(\lambda)-R(\lambda)D_q),
\end{aligned}\tag{2.7}$$

所以由 (2.7) 知 (F′) 的 (左端减右端) 的 Laplace 变换为

$$\begin{aligned}
&R(\lambda)-(\lambda I+D_q)^{-1}(R(\lambda)S+D_qR(\lambda)-R(\lambda)D_q+I)\\
=&(\lambda I+D_q)^{-1}(R(\lambda)(\lambda I-Q)-I),
\end{aligned}$$

所以由 Laplace 变换之唯一性知

$$P(t)=\Delta(t)+\int_0^t \Delta(t-s)(D_qP(s)+P(s)S-P(s)D_q)ds,\quad \text{a.s.}\tag{2.8}$$

的充分必要条件是

$$R(\lambda)(\lambda I-Q)=I\quad (\forall \lambda>0).\tag{2.9}$$

但是 (2.8) 左、右两边均关于 $t$ 连续。所以 "(F′) $\Leftrightarrow$ ($\mathrm{F}_\lambda$)". 定理 2.3 证毕.

**定理 2.4 ($Q$ 过程的存在性).** 任意给定一个转移强度矩阵 $Q$, 必存在一个 $Q$ 过程 $\overline{P}(t)$, 它满足 (B) 和 (F), 而且对任何 $Q$ 过程 $P(t)$ 来说, 恒有 $P(t)\geqslant \overline{P}(t)(\forall t\geqslant 0)$, 即 $\overline{P}(t)$ 是最小的 $Q$ 过程.

**证** 令 $D_q, S, \Delta(t)$ 之意义如定理 2.3. 显然 $S\geqslant 0$. 令

$$\begin{aligned}
&P_0(t)=\Delta(t),\\
&P_n(t)=\int_0^t \Delta(t-s)SP_{n-1}(s)ds\quad (n\geqslant 1),\\
&\overline{P}(t)=\sum_{n=0}^\infty P_n(t),
\end{aligned}\tag{2.10}$$

要证 $\overline{P}(t)$ 即为所求.

(1) 首先证明 $\overline{P}(t)$ 是一个标准的转移矩阵.

由 $S \geqslant 0$ 和 $\varDelta(t) \geqslant 0$ 知 $\overline{P}(t) \geqslant 0$. 而 $\varDelta(t)\mathbf{1} \leqslant \mathbf{1}$, 若设 $\sum\limits_{k=0}^{n} P_k(t)\mathbf{1} \leqslant \mathbf{1}$, 则

$$\begin{aligned}\sum_{k=0}^{n+1} P_k(t)\mathbf{1} &= P_0(t)\mathbf{1} + \int_0^t \varDelta(t-s)S\sum_{k=0}^{n} P_k(s)\mathbf{1}ds \\ &\leqslant P_0(t)\mathbf{1} + \int_0^t \varDelta(t-s)S\mathbf{1}ds.\end{aligned}$$

而 $Q\mathbf{1} \leqslant 0, S = Q + D_q$, 所以

$$\begin{aligned}\sum_{k=0}^{n+1} P_k(t)\mathbf{1} &\leqslant P_0(t)\mathbf{1} + \int_0^t \varDelta(t-s)D_q\mathbf{1}ds \\ &= \varDelta(t)\mathbf{1} + \mathbf{1} - \varDelta(t)\mathbf{1} = \mathbf{1}.\end{aligned}$$

由归纳法知: 对任何 $m \geqslant 0$ 有 $\sum\limits_{k=0}^{m} P_k(t)\mathbf{1} \leqslant \mathbf{1}$, 从而 $\overline{P}(t)\mathbf{1} \leqslant \mathbf{1}$.

下面证明 $\overline{P}(t)$ 满足 (K – C) 方程式. 先用归纳法证明:

$$P_n(s+t) = \sum_{k=0}^{n} P_k(s)P_{n-k}(t) \quad (\forall s, t \geqslant 0, n \geqslant 0). \tag{2.11}$$

由 $P_0(t) = \varDelta(t)$ 知 (2.11) 对 $n = 0$ 成立. 设 (2.11) 对 $n$ 成立. 则由 (2.10) 知

$$\begin{aligned}P_{n+1}(s+t) &= \int_0^{s+t} \varDelta(s+t-u)SP_n(u)du \\ &= \int_0^t \varDelta(s)\varDelta(t-u)SP_n(u)du + \int_0^s \varDelta(s-u)SP_n(t+u)du \\ &= \varDelta(s)\int_0^t \varDelta(t-u)SP_n(u)du + \int_0^s \varDelta(s-u)S\sum_{k=0}^{n} P_k(u)P_{n-k}(t)du \\ &= P_0(s)P_{n+1}(t) + \sum_{k=0}^{n}\left(\int_0^s \varDelta(s-u)SP_k(u)du\right)P_{n-k}(t) \\ &= P_0(s)P_{n+1}(t) + \sum_{k=0}^{n} P_{k+1}(s)P_{n-k}(t) \\ &= \sum_{k=0}^{n+1} P_k(s)P_{n+1-k}(t).\end{aligned}$$

(2.11) 证毕. 由 (2.11) 立即得

$$\overline{P}(s+t) = \overline{P}(s)\overline{P}(t) \quad (s, t \geqslant 0). \tag{2.12}$$

由

$$\sum_{k=0}^{n} P_k(t) = \varDelta(t) + \int_0^t \varDelta(t-s)S\sum_{k=0}^{n-1} P_k(s)ds$$

得

$$\overline{P}(t) = \Delta(t) + \int_0^t \Delta(t-s)S\overline{P}(s)ds. \tag{2.13}$$

所以由 $\Delta(t)$ 之定义得

$$\lim_{t\to 0+} \overline{P}(t) = \overline{P}(0) = I.$$

总之, $\overline{P}(t)$ 是标准的准转移矩阵.

(2) 其次, 证明 $\overline{P}(t)$ 是一个 $Q$ 过程. 事实上,

$$\frac{\overline{P}(t)-I}{t} = \frac{\Delta(t)-I}{t} + \frac{1}{t}\int_0^t \Delta(t-s)S\overline{P}(s)ds,$$

令 $t \to 0+$, 则上式趋于 $-D_q + S = Q$.

(3) 再次, 证明对任何 $Q$ 过程 $P(t)$, 都有 $P(t) \geqslant \overline{P}(t)$. 事实上, 若记 $P(t) = (p_{i,j}(t), i, j \in E)$, 则 $p_{i,i}(t) \geqslant e^{-q_i t}$, 所以 $P(t) \geqslant P_0(t)$. 设

$$P(t) \geqslant \sum_{k=0}^{n} P_k(t), \tag{2.14}$$

则由定理 2.2(1)$'$ 和 (2.14) 有:

$$\begin{aligned} P(t) &\geqslant \Delta(t) + \int_0^t \Delta(t-s)S\sum_{k=0}^{n} P_k(s)ds \\ &= \sum_{k=0}^{n+1} P_k(t). \end{aligned}$$

所以 $P(t) \geqslant \sum\limits_{k=0}^{m} P_k(t)(\forall m \geqslant 0)$, 从而

$$P(t) \geqslant \lim_{m\to\infty} \sum_{k=0}^{m} P_k(t) = \overline{P}(t).$$

(4) 最后, 证明 $\overline{P}(t)$ 满足 (B) 和 (F).

事实上, 由 (2.13) 及定理 2.2 知 $\overline{P}(t)$ 满足 (B). 至于 (F), 令 $\overline{P}(t)$ 和 $P_n(t)$ 的 Laplace 变换分别为 $\overline{R}(\lambda)$ 和 $R_n(\lambda)$. 容易算出: $R_0(\lambda) = (\lambda I + D_q)^{-1}, R_n(\lambda) = \Pi(\lambda)R_{n-1}(\lambda)(n \geqslant 1)$, 其中 $\Pi(\lambda) = (\lambda I + D_q)^{-1}S$. 所以

$$\begin{aligned} \overline{R}(\lambda) &= \sum_{n=0}^{\infty} R_n(\lambda) = \sum_{n=0}^{\infty} \Pi(\lambda)^n (\lambda I + D_q)^{-1} \\ &= (\lambda I + D_q)^{-1} + \sum_{n=0}^{\infty} \Pi(\lambda)^n (\lambda I + D_q)^{-1} S (\lambda I + D_q)^{-1}, \end{aligned}$$

从而

$$\overline{R}(\lambda)(\lambda I + D_q) = I + \overline{R}(\lambda)S,$$

即是 $\overline{R}(\lambda)(\lambda I - Q) = I$, 亦即 $\overline{P}(t)$ 满足 (F). 至此, 定理证毕.

下面我们研究 $Q$ 过程唯一的充分必要条件. 这里只给出对保守的转移强度矩阵 $Q$ 来说, $Q$ 过程唯一的充分必要条件. 对一般的转移强度矩阵 $Q$, 其 $Q$ 过程的充分必要条件请参看 [34].

**定理 2.5** 给定保守的转移强度矩阵 $Q$, 令 $\overline{R}(\lambda), R_n(\lambda), S, D_q, \Pi(\lambda)$ 如定理 2.4 中所定义, 记 $S_n(\lambda) \stackrel{\text{def.}}{=\!=} \sum\limits_{k=0}^{n} R_k(\lambda), \overline{y}(\lambda) = \mathbf{1} - \lambda\overline{R}(\lambda)\mathbf{1}$, 则

(1) 当 $n \uparrow \infty$ 时, $\Pi^n(\lambda)\mathbf{1}$ 单调下降到 $\overline{y}(\lambda)$;

(2) $\overline{y}(\lambda)$ 是《$\Pi(\lambda)y = y, 0 \leqslant y \leqslant \mathbf{1}$》的最大解;

(3) $\overline{y}(\lambda) = 0$ 的充分必要条件是

《$(\lambda I - Q)y = 0, 0 \leqslant y \leqslant \mathbf{1}$》仅有零解. 此处 $y, \overline{y}(\lambda)$ 都是列向量, 其维数与状态空间 $E$ 的维数一样.

**证** (1) 由于

$$S_n(\lambda) = R_0(\lambda) + (I + \Pi(\lambda) + \cdots + \Pi(\lambda)^{n-1})(\lambda I + D_q)^{-1} \cdot S(\lambda I + D_q)^{-1},$$

所以

$$S_n(\lambda)(\lambda I + D_q) = I + S_{n-1}(\lambda)S.$$

若注意 $Q\mathbf{1} = 0, Q = S - D_q$, 则得

$$\lambda S_n(\lambda)\mathbf{1} + S_n(\lambda)D_q\mathbf{1} = \mathbf{1} + S_{n-1}(\lambda)S\mathbf{1} = \mathbf{1} + S_{n-1}(\lambda)D_q\mathbf{1}. \tag{2.15}$$

故

$$S_n(\lambda)D_q\mathbf{1} \leqslant \mathbf{1} + S_{n-1}(\lambda)D_q\mathbf{1}.$$

而

$$S_0(\lambda)D_q\mathbf{1} = R_0(\lambda)D_q\mathbf{1} = (\lambda I + D_q)^{-1}D_q\mathbf{1} < \infty,$$

所以, 对一切 $n \geqslant 0$ 有:

$$S_n(\lambda)D_q\mathbf{1} < \infty.$$

所以可以在 (2.15) 左、右两边减去 $S_{n-1}(\lambda)D_q\mathbf{1}$, 减去后得

$$\lambda S_n(\lambda)\mathbf{1} + R_n(\lambda)D_q\mathbf{1} = \mathbf{1},$$

即

$$\lambda S_n(\lambda)\mathbf{1} + \Pi(\lambda)^n(\lambda I + D_q)^{-1}D_q\mathbf{1} = \mathbf{1}.$$

注意 $S\mathbf{1} = D_q\mathbf{1}$, 则由上式可得:

$$\lambda S_n(\lambda)\mathbf{1} + \Pi(\lambda)^{n+1}(\lambda)\mathbf{1} = \mathbf{1},$$

即 $\Pi(\lambda)^{n+1}\mathbf{1} = \mathbf{1} - \lambda S_n(\lambda)\mathbf{1}$.
但是, 当 $n \uparrow \infty$ 时 $S_n(\lambda) \uparrow \overline{R}(\lambda)$, 所以

$$\Pi(\lambda)^n\mathbf{1} \downarrow (1 - \lambda\overline{R}(\lambda)\mathbf{1}) \quad (\text{当 } n \uparrow \infty).$$

此即

$$\Pi(\lambda)^n\mathbf{1} \downarrow \overline{y}(\lambda) \quad (\text{当 } n \uparrow \infty).$$

(2) 由 (1) 有

$$\begin{aligned}\Pi(\lambda)\overline{y}(\lambda) &= \Pi(\lambda) \lim_{n\to\infty} \Pi(\lambda)^n\mathbf{1} \\ &= \lim_{n\to\infty} \Pi(\lambda)^{n+1}\mathbf{1} = \overline{y}(\lambda).\end{aligned}$$

显然

$$\overline{y}(\lambda) = \lambda \int_0^\infty e^{-\lambda t}(\mathbf{1} - P(t)\mathbf{1})dt,$$

所以

$$0 \leqslant \overline{y}(\lambda) \leqslant \mathbf{1}.$$

这就证明了: $\overline{y}(\lambda)$ 是《$\Pi(\lambda)y = y, 0 \leqslant y \leqslant \mathbf{1}$》的一个解. 今设 $y$ 为另一解, 则 $\Pi(\lambda)y = y, 0 \leqslant y \leqslant \mathbf{1}$, 故 $y = \Pi(\lambda)^n y \leqslant \Pi(\lambda)^n\mathbf{1}$, 令 $n \to \infty$ 即得 $y \leqslant \overline{y}(\lambda)$. 这就证明了 $\overline{y}(\lambda)$ 是最大解.

(3) 由 (2) 得知: $\overline{y}(\lambda) = 0$ 的充分必要条件是《$\Pi(\lambda)y = y, 0 \leqslant y \leqslant \mathbf{1}$》只有零解.

而 $\Pi(\lambda) = (\lambda I + D_q)^{-1}S$, 所以, 当 $0 \leqslant y \leqslant \mathbf{1}$ 时, 有:

$$\Pi(\lambda)y = y \Leftrightarrow Sy = (\lambda I + D_q)y \Leftrightarrow (\lambda I - Q)y = 0.$$

这就证明了 (3). 定理 2.5 证毕.

**定理 2.6** 给定保守的转移强度矩阵 $Q$, 恰有唯一一个 $Q$ 过程的充分必要条件是:

$$《(\lambda I - Q)y = 0, \quad 0 \leqslant y \leqslant \mathbf{1}》$$

只有零解.

**证** 由定理 2.5 即得定理 2.6.

**定义 2.4** 称转移强度矩阵 $Q$ 是**正则的**, 如果其最小 $Q$ 矩阵 $\overline{P}(t)$ 是不间断的, 即 $\overline{P}(t)\mathbf{1} = \mathbf{1}(\forall t \geqslant 0)$.

**定理 2.7** 设 $Q$ 是任一转移强度矩阵, 则有:

(1) $Q$ 是正则的 $\Rightarrow Q$ 过程唯一;

(2) $Q$ 是正则的 $\Leftrightarrow \overline{y}(\lambda) = 0$;

(3) $Q$ 是正则的 $\Rightarrow Q$ 是保守的.

证明甚易, 读者可作为习题验证之.

**思考题** 当 $Q$ 过程不唯一时, 如何构造全部 $Q$ 过程? 当精确的 $Q$ 过程不易求时, 能否有渐近 $Q$ 过程?

## §3 转移矩阵之遍历性及遍历矩阵之性质

在第六章离散时间可数状态时齐的 Markov 链中, 其转移矩阵 $P = (p_{i,j}, i, j \in E)$, 极限

$$\lim_{n\to\infty} P^n \tag{3.1}$$

未必存在, 但极限

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{1}{n} P^i = \Pi \tag{3.2}$$

一定存在.

第六章定理 5.4 证明了极限 (3.1) 存在的充分必要条件是每个正状态的周期为 1. (注意: 由 (K − C) 方程式, 哪里的 $P^{(n)} = P^n$.) 第六章 §5 自定理 5.4 以后主要研究极限 (3.2) 的性质及求法.

现在第七章研究的是: 连续时间的可数状态的时齐的 Markov 过程, 其转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$, 我们假定了它是标准的. 在此假定下, 定理 1.2 证明了极限

$$\lim_{t\to\infty} P(t) = \Pi \tag{3.3}$$

永远存在.

对标准的准转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$ 而言, 用扩展矩阵的方法, 取 $\Delta \overline{\in} E$, 令

$$P_\Delta(t) = \begin{array}{c} \\ E \\ \Delta \end{array}\!\!\begin{array}{c} \begin{array}{cc} E & \quad\Delta \end{array} \\ \begin{pmatrix} P(t) & d(t) \\ 0 & 1 \end{pmatrix} \end{array}, \quad d(t) = \mathbf{1} - P(t)\mathbf{1},$$

则极限

$$\lim_{t\to\infty} P_\Delta(t) = \Pi_\Delta \tag{3.4}$$

恒存在, 而 $P(t)$ 是 $P_\Delta(t)$ 在 $E\times E$ 上的局限, 所以极限

$$\lim_{t\to\infty} P(t)=\Pi \tag{3.5}$$

恒存在. 在这一节中, 我们将要研究 $\Pi$ 的性质及求法. 称 $\Pi$ 是 $P(t)$ 的**遍历矩阵**.

在这一节中, 恒记

$$(l)=\left\{(a_i, i\in E):\sum_{i\in E}|a_i|<\infty\right\},$$

$$(m)=\left\{\begin{pmatrix} a_i \\ i\in E\end{pmatrix}:\sup_{i\in E}|a_i|<\infty\right\}.$$

$(l)$ 中的元素 (行向量) 用 $\alpha',\beta',\gamma',\cdots$ 表示, $(m)$ 中的元素 (列向量) 用 $x,y,z,\cdots$ 表示.

对任何转移强度矩阵 $Q$, 仍用 $\overline{P}(t)$ 表示其最小 $Q$ 过程, 其构造见定理 2.4, $\overline{\Pi}$ 是 $\overline{P}(t)$ 的遍历矩阵.

本节沿袭 §2 的符号.

**定理 3.1** 设 $Q$ 是正则的转移强度矩阵, $\alpha'\in(l),\alpha'\geqslant 0$, 则

$$\alpha' Q=0\Leftrightarrow \alpha'\overline{\Pi}=\alpha'.$$

**证** 设 $\alpha'\in(l),\alpha'\geqslant 0,\alpha'\overline{\Pi}=\alpha'$. 由定理 1.2 (该定理对标准的准转移矩阵也成立) 有

$$\alpha'=\alpha'\overline{\Pi}=\alpha'\overline{\Pi}\,\overline{P}(t)=\alpha'\overline{P}(t)\quad(\forall t\geqslant 0).$$

把上式左右两边取 Laplace 变换得:

$$\alpha'=\lambda\alpha'\overline{R}(\lambda)\quad(\lambda>0). \tag{3.6}$$

由 $\overline{R}(\lambda)$ 满足 $(\mathrm{F}_\lambda)$ 得

$$\overline{R}(\lambda)(\lambda I+D_q)=\overline{R}(\lambda)S+I\quad(\lambda>0). \tag{3.7}$$

把 (3.7) 两边左乘以 $\alpha'$, 并注意 (3.6) 得

$$\frac{1}{\lambda}\alpha' D_q=\frac{1}{\lambda}\alpha' S\quad(\lambda>0).$$

此即 $\alpha' Q=0$.

再设 $\alpha'\in(l),\alpha'\geqslant 0,\alpha' Q=0$, 则

$$\alpha'(\lambda I-Q)=\lambda\alpha'\geqslant 0(\forall\lambda>0).$$

所以

$$\alpha'(\lambda I+D_q)=\lambda\alpha'+\alpha' S.$$

而

$$\alpha' S=\alpha'(\lambda I+D_q)\Pi(\lambda),$$

反复利用上述两式得

$$\alpha'(\lambda I+D_q)\geqslant\sum_{k=0}^{n}(\lambda\alpha')\Pi(\lambda)^k\quad(n\geqslant 1),$$

即

$$\begin{aligned}\alpha'&\geqslant\sum_{k=0}^{n}(\lambda\alpha')\Pi(\lambda)^k(\lambda I+D_q)^{-1}\\&=\sum_{k=0}^{n}(\lambda\alpha')R_k(\lambda)\quad(\forall n\geqslant 0,\lambda>0),\end{aligned}\tag{3.8}$$

($\Pi(\lambda)$ 和 $R_k(\lambda)$ 之定义见定理 2.4.) 在 (3.8) 中对 $n\to\infty$ 取极限得

$$\alpha'\geqslant\lambda\alpha'\overline{R}(\lambda)\quad(\forall\lambda>0).\tag{3.9}$$

由于 $Q$ 正则, 所以 $\lambda\overline{R}(\lambda)\mathbf{1}=\mathbf{1}$, 从而 (3.9) 只能两边相等. 但是

$$\lim_{\lambda\to 0+}\lambda\overline{R}(\lambda)=\lim_{t\to\infty}\overline{P}(t)=\overline{\Pi},\tag{3.10}$$

$\lambda\overline{R}(\lambda)\geqslant 0,\lambda\overline{R}(\lambda)\mathbf{1}=\mathbf{1},\alpha'\in(l),\alpha'\geqslant 0$, 所以, 在 (3.9) 中令 $\lambda\to 0+$, 并用控制收敛定理可得

$$\alpha'=\alpha'\overline{\Pi}.$$

定理证毕.

**定理 3.2** 设 $Q$ 是任一转移强度矩阵, 则

$$\overline{\Pi}Q=Q\overline{\Pi}=0.$$

**证** 因为 $\overline{P}'(t)=Q\overline{P}(t),\lim\limits_{t\to\infty}\overline{P}(t)=\overline{\Pi}$, 而且 $\lim\limits_{t\to\infty}\overline{P}'(t)=0$, 所以用控制收敛定理得

$$\begin{aligned}0&=\lim_{t\to\infty}\overline{P}'(t)=\lim_{t\to\infty}Q\overline{P}(t)\\&=Q\left(\lim_{t\to\infty}\overline{P}(t)\right)=Q\overline{\Pi}.\end{aligned}$$

又因为 $\overline{\Pi}=\overline{\Pi}^2$, 所以由定理 3.1 (注意: 定理 3.1 的充分性部分不需 $Q$ 是 "正则" 这一条件, 对一切转移强度都对) 知: $\overline{\Pi}Q=0$. 定理证毕.

下面我们将要研究遍历矩阵 $\Pi$ 的结构, 并给出如何由转移强度矩阵 $Q$ 来求其 $Q$ 过程的遍历矩阵 $\Pi$ 的方法. 由于在实际问题中, 只知道转移强度矩阵 $Q$ 而不知道其对应的 $Q$ 过程 $P(t)$, 但是, 知道 $P(t)$ 的遍历极限 $\Pi = \lim\limits_{t\to\infty} P(t)$ 就能解决问题. 所以, 直接通过 $Q$ 求出 $\Pi$, 在实际中往往是很有用的.

**定理 3.3** 任给转移强度矩阵 $Q$, 其 $Q$ 过程的遍历极限 $\Pi = \lim\limits_{t\to\infty} P(t)$ 具有下述结构:

$$\Pi = \begin{array}{c} \\ N \\ R_1 \\ R_2 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} N & R_1 & R_2 & \cdots \end{array} \\ \begin{pmatrix} 0 & a(1)\pi'(1) & a(2)\pi'(2) & \cdots \\ 0 & \mathbf{1}\pi'(1) & 0 & \cdots \\ 0 & 0 & \mathbf{1}\pi'(2) & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \end{array}, \tag{3.11}$$

其中 $E = N\bigcup R = N\bigcup\left(\bigcup\limits_n R_n\right)$, $N$ 或 $R$ 可以是空集, $N\bigcap R = \varnothing, R_n\bigcap R_m = \varnothing$ (当 $m\neq n$), $a(k)\geqslant 0$ 是列向量, 其维数与 $N$ 中的元素的个数相同, $\pi'(k)$ 是行向量, 其维数与 $R_k$ 中的元素的个数相同, $\pi'(k)\geqslant 0, \pi'(k)\mathbf{1} = \mathbf{1}, \Pi\mathbf{1}\leqslant\mathbf{1}$. 注意: $\mathbf{1}\pi'(k)$ 是行行都是 $\pi'(k)$ 的转移矩阵.

**证** 由于

$$\lim_{t\to\infty} P(t) = \Pi$$

恒存在, 故

$$\lim_{n\to\infty} P(1)^n = \Pi$$

恒存在, 而 $P(1)$ 是以实数为元素的准转移矩阵, 所以由第六章定理 5.5 即得本定理.

注意定理 3.3 中的 $N$ 相当于第六章定理 5.5 中的 $N\bigcup R^0, R_n$ 相当于第六章定理 5.5 中的 $R_n^+$. 由于 $\Pi$ 对应于 $N\bigcup R^0$ 中的元素的列都是 0, 所以我们把第六章定理 5.5 中的 $\Pi$ 的结构写成定理 3.3 中的简单形式.

此处 $\pi'(k)$ 是以实数为元素的行向量, 不是函数的微商.

设 $Q$ 是正则的转移强度矩阵. 由定理 3.1 知, 当 $\alpha'\in(l), \alpha'\geqslant 0$ 时,

$$\alpha' Q = 0 \Leftrightarrow \alpha'\overline{\Pi} = \alpha'.$$

若再注意 $\overline{\Pi}$ 有定理 3.3 中的结构 (3.11), 可以看出:

$$《\alpha' Q = 0,\ \ \alpha'\in(l),\ \ \alpha'\geqslant 0》$$

的通解必为

$$\begin{matrix} & N & R_1 & R_2 & \\ \alpha' = ( & 0, & c_1\pi'(1), & c_2\pi'(2), & \cdots) \end{matrix}$$

其中 $c_i$ 是非负实数且 $\sum_i c_i < \infty$. 所以, 解方程式

$$《\alpha' Q = 0, \quad \alpha' \in (l), \quad \alpha' \geqslant 0》$$

就可以判断: 其通解中那些必为 0 的分量所对应的位置 (状态) 就是 $N$, 也可以看出 $R_i$ 及对应的 $\pi'(i)$. 所以, 若再把 $a(i)$ 这些列向量算出来, 则 $\overline{\Pi}$ 就完全算出来了. 下面就来研究 $a(i)$ 如何求.

**定理 3.4** 设 $Q$ 是任意一个转移强度矩阵, 则

(1) $y \geqslant 0, y \in (m), Qy = 0 \Rightarrow y \geqslant \overline{\Pi} y$;

(2) $a(1)$ 是方程式

$$(Z): 《0 = Q \begin{pmatrix} Z \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix} \begin{matrix} N \\ R_1 \\ R_2 \\ \vdots \end{matrix}, \quad Z \geqslant 0, Z \in (m)》$$

的最小解 $y_{\min}$ 中所对应的 $Z$, 即 $a(1) = y_{\min}$ on $N$. (其中 $a(i)$ 的求法类似.)

**证** (1) 因为 $Qy = 0, y \geqslant 0, y \in (m)$, 所以

$$(\lambda I - Q)y = \lambda y \quad (\forall \lambda > 0).$$

因此, 仿 (3.9) 的证法有 (注意 (3.9) 式的证明并不要求 $Q$ 是正则的):

$$y \geqslant \lambda \overline{R}(\lambda) y. \tag{3.12}$$

在 (3.12) 中令 $\lambda \to \infty$, 并注意 $y \geqslant 0$ 和 $y \in (m)$, 再用 Fatou 引理可得

$$\begin{aligned} y &\geqslant \liminf_{\lambda \to 0+} \lambda \overline{R}(\lambda) y \geqslant \left( \liminf_{\lambda \to 0+} \lambda \overline{R}(\lambda) \right) y \\ &= \left( \liminf_{t \to \infty} \overline{P}(t) \right) y = \overline{\Pi} y. \end{aligned} \tag{3.13}$$

(2) 由定理 3.2 有: $Q\overline{\Pi} = 0$, 更有 $Q\overline{\Pi} e_j = 0 (\forall j \in R_1$, 其中 $e_j$ 是第 $j$ 个坐标轴上的单位列向量, 即 $e_j$ on $\{j\} = 1, e_j$ on $(E - \{j\})$ 为零列向量), 此即

$$Q \begin{pmatrix} a(1) \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix} = 0.$$

显然 $0 \leqslant a(1) \in (m)$, 所以

$$y = \begin{matrix} N \\ R_1 \\ R_2 \\ \vdots \end{matrix} \begin{pmatrix} a(1) \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix}$$

是方程式 $(Z)$ 的解. 再设

$$\begin{pmatrix} Z \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix}$$

也是方程式 $(Z)$ 的解, 则

$$\begin{pmatrix} Z \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix} \geqslant \overline{\Pi} \begin{pmatrix} Z \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix} \geqslant \overline{\Pi} \begin{pmatrix} 0 \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix} = \begin{pmatrix} a(1) \\ \mathbf{1} \\ 0 \\ \vdots \end{pmatrix}.$$

定理证毕.

下面我们研究一类简单的正则转移强度矩阵 $Q$ 所对应的 $\overline{\Pi}$ 的求法.

**定义 3.1** *称转移强度矩阵* $Q=(q_{i,j}, i, j \in E)$ *是***可约的**, *如果存在* $E$ *的一个真子集* $F \subset E, F \neq E$, *使得对任何* $i \in F, j \overline{\in} F$, *都有* $q_{i,j}=0$. *反之称* $Q$ *是***不可约的**.

**定理 3.5** *若* $Q$ *是不可约的转移强度矩阵, 则有: 或者* $E=N$ (*这时* $\overline{\Pi}=0$), *或者* $E=R_1$ (*这时* $\overline{\Pi}=\mathbf{1}\pi'(1)$).

**证**　若 $E=N \bigcup \left(\bigcup\limits_n R_n\right)$ 中有 $N \neq \varnothing$, 且 $R_1 \neq \varnothing$, 则 $\overline{\Pi}$ 中必有一行 $e_i'\overline{\Pi}$ ($e_i'$ 是第 $i$ 个轴上之分量为 1, 其他轴上之分量为 0 的行向量) 具有下述形式:

$$\begin{matrix} R_1 & N \bigcup \left( \bigcup\limits_{n>1} R_n \right) \\ (\pi'(1) & 0 \qquad ). \end{matrix}$$

记 $\overline{P}(t)$ 如下:

$$\begin{array}{cc} & \begin{array}{cc} R_1 & N\cup\left(\bigcup\limits_{n>1}R_n\right) \end{array} \\ \overline{P}(t)= \begin{array}{c} R_1 \\ N\cup\left(\bigcup\limits_{n>1}R_n\right) \end{array} & \begin{pmatrix} A(t) & B(t) \\ C(t) & D(t) \end{pmatrix} \end{array}$$

在 $\overline{\Pi}\ \overline{P}(t)=\overline{\Pi}$ 中取前面指出的那一行:

$$e_i'\overline{\Pi}\ \overline{P}(t)=e_i'\overline{\Pi},$$

此即

$$(\pi'(1),0)\begin{pmatrix} A(t) & B(t) \\ C(t) & D(t) \end{pmatrix}=(\pi'(1),0).$$

所以

$$\pi'(1)B(t)=0.$$

但是 $\pi'(1)>0$, 所以 $B(t)=0$. 故

$$Q=\begin{pmatrix} A'(0) & 0 \\ C'(0) & D'(0) \end{pmatrix},$$

即是 $Q$ 是可约的. 定理证毕.

**定理 3.6**　设 $Q$ 是不可约的正则的转移强度矩阵, 则

(1)《$\alpha'Q=0,\alpha'\geqslant 0,\alpha'\in(l)$》只有零解 $\Rightarrow \overline{\Pi}=0$;

(2)《$\alpha'Q=0,\alpha'\geqslant 0,\alpha'\in(l)$》有非零解 $\Rightarrow$ 其通解必为 $c\pi',\pi'\geqslant 0,\pi'\mathbf{1}=1$, 这时 $\overline{\Pi}=\mathbf{1}\pi'$.

**证**　由定理 3.3、定理 3.5 即得定理 3.6.

## §4　分枝过程与种群繁衍

本节主要讲实际问题与 Markov 过程的应用. 有些证明很繁的定理就略去了证明, 但标明其出处. 有兴趣的读者可查阅.

自然界的某一物种的繁衍、原子反应堆中质点的裂变、人体中某类细胞的分裂等现象中, 其数学模型常常涉及一类很重要的随机过程 —— 分枝过程. 这类过程是怎样提炼出来的呢? 它的概率特征是什么? 我们主要关心的是什么问题? 下面通过一些例子来剖析这些问题.

**例 4.1 (细胞分裂)** 考虑人体中某种细胞的分裂现象.

设在时刻 0 人体中某种细胞的个数为 $X_0$, 经过一个单位时间以后, 这 $X_0$ 个细胞中的每一个可分裂成 0 个 (即死亡)、1 个 (即不分裂), $\cdots, m$ 个等等. 由于营养、药物刺激、运动、疲劳等因素的影响, 一个细胞分裂成 $m$ 个是有确定的概率 $p_m$ 的, 其中 $p_m \geqslant 0, \sum\limits_{m=0}^{\infty} p_m = 1$. 一般可以假定各个不同的细胞的分裂结果是相互独立的, 而且具有公共分布 $\{p_m : m \geqslant 0\}$. 于是

$$X_1 = \sum_{i=1}^{X_0} \xi_i^{(1)},$$

仿之

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_i^{(n)} \quad (n \geqslant 1),$$

其中 $\{\xi_i^{(n)} : n = 1, 2, \cdots, i = 1, 2, \cdots\}$ 相互独立且具有公共分布.

**例 4.2 (反应堆中的质点的裂变)** 设在时刻 0 某反应堆中某类质点有 $X_0$ 个, 由于质点之间的相互碰撞或其他射线的轰击, 每隔一个单位时间, 一个质点可以分裂成 $m$ 个质点 $(m = 0, 1, 2, \cdots)$, 其对应的概率为 $p_m\left(p_m \geqslant 0, \sum\limits_{m=0}^{\infty} p_m = 1\right)$. 一般而言, 仍可假定:

(1) 各质点的分裂结果是相互独立的, 具有公共分布, 即每一质点经过一个单位时间分裂成的新质点数是一个随机变量, 且这些随机变量是独立同分布的.

(2) 质点的分裂情况与 "年龄" 无关, 如果用 $\xi_i^{(n)}$ 代表在时刻 $n-1$ 存在的第 $i$ 个质点经过一个单位时间后分裂成的质点数, 则 $\{\xi_i^{(n)} : i = 1, 2, \cdots, n = 1, 2, \cdots\}$ 是独立同分布的, 且

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_i^{(n)},$$

其中 $X_n$ 是时刻 $n$ 该类质点的个数.

从例 4.1 和 4.2 可以看出: 尽管这两类例子的实际背景不同, 但它们的数学模型是一样的. 其实, 还有很多实际问题, 它们都属于这一数学模型.

下面我们给出分枝过程的数学定义.

**定义 4.1** 设 $\{\xi_i^{(n)} : i = 1, 2, \cdots, n = 1, 2, \cdots\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一族独立同分布的取非负整数的随机变量, 其公共分布为 $\{p_m : m \geqslant 0\}$. 若 $X_0$ 是 $(\Omega, \mathscr{F}, P)$ 上的任一取正整数的随机变量, 且

$$X_n = \sum_{i=1}^{X_{n-1}} \xi_i^{(n)} \quad (n \geqslant 1), \tag{4.1}$$

则称 $\{X_n : n = 0, 1, \cdots\}$ 是一个离散时间的**分枝过程**或**分枝链**. 这就是所谓的 **Galton-Watson 分枝过程**.

令 $f(s) = \sum\limits_{m=0}^{\infty} p_m s^m (|s| \leqslant 1)$, 称 $f$ 为分布 $\{p_m\}$ 的母函数, 有时也称 $f$ 为 $\xi_i^{(n)}$ 的母函数, 或 $\{X_n\}$ 的本原母函数.

显然 Galton-Watson 分枝过程 $\{X_n\}$ 是一时齐的可数状态的 Markov 链. 若令

$$p_{i,j}^{(k)} = P(X_{n+k} = j | X_n = i) \tag{4.2}$$

是其 $k$ 步转移概率, 令

$$P^{(k)} = (p_{i,j}^{(k)}, i, j \geqslant 0) \tag{4.3}$$

是其 $k$ 步 (阶) 转移矩阵, 则有

$$\begin{cases} p_{i,j}^{(k)} = \sum\limits_{\sum\limits_{s=1}^{i} j_s = j} \prod\limits_{s=1}^{i} p_{1,j_s}^{(k)}, & \text{当 } i > 0, j, k \geqslant 0, \\ p_{0,j}^{(k)} = \delta_{0,j}, & \text{当 } j, k \geqslant 0, \end{cases} \tag{4.4}$$

其中 $\delta_{i,j} = 0$ 或 $1$ 视 $i \neq j$ 或 $i = j$ 而定.

显然 $p_{1,m} \stackrel{\text{def.}}{=\!=} p_{1,m}^{(1)} = p_m$, 而且对任何 $i \geqslant 0, k \geqslant 1, \{p_{i,m}^{(k)} : m \geqslant 0\}$ 是一个概率分布. 令 $g_{k,i}(s) = \sum\limits_{m=0}^{\infty} p_{i,m}^{(k)} s^m$ 为其母函数.

对于分枝链 $\{X_n\}$, 研究它的矩、分布、极限分布, 特别是极限分布中的两个特例

$$P\left(\lim_{n\to\infty} X_n = \infty\right) \text{ 与 } P\left(\lim_{n\to\infty} X_n = 0\right),$$

有着重要的意义. 前者意味着该系统要 "爆炸" 的概率, 而后者是该系统要 "熄灭" 的概率. 如果 $X_n$ 代表某一患者在时刻 $n$ 癌细胞的个数, 前者就是 "病情急剧恶化" 的概率, 而后者正是该患者终将 "治愈" 的概率.

对分枝链 $\{X_n\}$ 的研究, 母函数 $f(s)$ 扮演着一个重要角色.

**定理 4.1** 设 $\{X_n : n \geqslant 0\}$ 是分枝链. 若 $X_0 \equiv n_0$ 是一正整数, $f_n(s)$ 是 $X_n$ 的矩母函数, $\mu_n = \boldsymbol{E}(X_n), \sigma_n^2 = \operatorname{var}(X_n)(n \geqslant 0), f(s)$ 是 $\{X_n\}$ 的本原矩母函数, 则

(1) $f_1(s) = f(s)^{n_0}$;

(2) $f_{n+1}(s) = f_n(f(s)), n \geqslant 0$;

(3) $\mu_n = n_0 \mu^n (n \geqslant 1, \mu = f'(1))$;

(4) $\sigma_n^2 = \begin{cases} \dfrac{n_0 \sigma^2 \mu^n (\mu^n - 1)}{\mu^2 - \mu}, & \text{当 } \mu \neq 1, \\ n_0 n \sigma^2, & \text{当 } \mu = 1 \end{cases} (n \geqslant 1),$

其中 $\sigma^2 = f''(1) + f'(1) - (f'(1))^2$.

**证** 通过直接计算立即可得 (1),(2),(3). 下面证明 (4). 事实上,

$$\sigma_n^2 = \mathrm{var}(X_n) = f_n''(1) + f_n'(1) - (f_n'(1))^2. \tag{4.5}$$

而由 (2) 有

$$\begin{aligned} f_n''(s) &= (f_{n-1}(f(s)))'' \\ &= f_{n-1}''(f(s))(f'(s))^2 + f''(s)f_{n-1}'(f(s)). \end{aligned} \tag{4.6}$$

所以若令

$$\alpha = n_0(\sigma^2 - \mu + \mu^2), \tag{4.7}$$

并注意:

$$\mu = f'(1), \quad f''(1) = \sigma^2 - \mu + \mu^2, \tag{4.8}$$

则由 (3) 及 (4.6),(4.7),(4.8) 得

$$\begin{aligned} f_n''(1) &= f_{n-1}''(1)\mu^2 + (\sigma^2 - \mu + \mu^2)f_{n-1}'(1) \\ &= f_{n-1}''(1)\mu^2 + \alpha\mu^{n-1} \quad (n \geqslant 2). \end{aligned} \tag{4.9}$$

反复利用 (4.9) 递推可得

$$f_n''(1) = f_1''(1)\mu^{2(n-1)} + \alpha(\mu^{2n-3} + \mu^{2n-4} + \cdots + \mu^{n-1}). \tag{4.10}$$

由 (4.10),(4.8) 可得

$$\begin{aligned} \sigma_n^2 &= \mathrm{var}(X_n) \\ &= f_n''(1) + f_n'(1) - (f_n'(1))^2 \\ &= f_1''(1)\mu^{2(n-1)} + \alpha(\mu^{2n-3} + \mu^{2n-4} + \cdots + \mu^{n-1}) + n_0\mu^n - (n_0\mu^n)^2 \\ &= [n_0(n_0-1)\mu^2 + \alpha]\mu^{2n-2} + \alpha(\mu^{2n-3} + \mu^{2n-4} + \cdots + \mu^{n-1}) \\ &\quad + n_0\mu^n - (n_0\mu^n)^2 \\ &= n_0(\mu^n - \mu^{2n}) + \alpha(\mu^{2n-2} + \mu^{2n-3} + \cdots + \mu^{n-1}). \end{aligned} \tag{4.11}$$

当 $\mu \neq 1$ 时, 由 (4.11) 得

$$\sigma_n^2 = \frac{n_0\sigma^2\mu^n(\mu^n - 1)}{\mu^2 - \mu}.$$

当 $\mu = 1$ 时, 由 (4.11) 得

$$\sigma_n^2 = n_0 n\sigma^2.$$

定理 4.1 得证.

**定义 4.2** 设 $\{X_n : n \geqslant 0\}$ 是一个分枝链, 称

$$\rho \stackrel{\text{def.}}{=\!=} \lim_{n\to\infty} P(X_n = 0) \tag{4.12}$$

为 $\{X_n\}$ 的**绝灭概率**.

由于 $\{X_n = 0\} \subset \{X_{n+1} = 0\}$, 所以

$$\begin{aligned}\rho &= \lim_{n\to\infty} P(X_n = 0) = P\left(\lim_{n\to\infty}\{X_n = 0\}\right)\\ &= P\left(\lim_{n\to\infty} X_n = 0\right).\end{aligned} \tag{4.13}$$

**定理 4.2** 设 $\{X_n : n = 0, 1, \cdots\}$ 是分枝链, $X_0 \equiv 1, f(s)$ 是其本原母函数, $\mu = f'(1), \rho$ 是其绝灭概率, $f(s) = \sum\limits_{m=0}^{\infty} p_m s^m$ 则

(1) $\rho = f(\rho)$;

(2) 当 $\mu \leqslant 1$ 且 $p_1 < 1$ 时, 有 $\rho = 1$;

(3) 当 $\infty \geqslant \mu > 1$ 时, $\rho$ 是方程式

$$《\, f(s) = s, 0 \leqslant s < 1 \,》 \tag{4.14}$$

的唯一解.

**证** 首先注意: 若令 $f_n(s)$ 是 $X_n$ 的母函数, 由于 $X_0 \equiv 1$, 所以 $f_1(s) = f(s)$. (1) 由 $f_{n+1}(s) = f_n(f(s)) = f(f_n(s))$ 及 $\rho$ 之定义即得

$$\begin{aligned}\rho &= \lim_{n\to\infty} P(X_n = 0) = \lim_{n\to\infty} f_n(0)\\ &= \lim_{n\to\infty} f(f_{n-1}(0)) = f\left(\lim_{n\to\infty} f_{n-1}(0)\right)\\ &= f(\rho).\end{aligned}$$

(2) (A) 先设 $\mu \leqslant 1, p_0 + p_1 < 1$, 则

$$f'(s) = \sum_{n=1}^{\infty} n p_n s^{n-1}$$

在 $[0,1)$ 上是严格上升函数. 但是 $\mu = f'(1) = \lim\limits_{s\uparrow 1} f'(s)$, 所以

$$\mu > f'(s) \quad (\forall s \in [0,1)). \tag{4.15}$$

假设 $\rho < 1$, 则存在 $c \in (\rho, 1)$ , 使

$$f'(c) = \frac{f(1) - f(\rho)}{1-\rho} = \frac{1-\rho}{1-\rho} = 1. \tag{4.16}$$

由 (4.15),(4.16) 得知

$$\mu > f'(c) = 1.$$

这与 $\mu \leqslant 1$ 矛盾, 所以 $\rho = 1$.

(B) 再设 $\mu \leqslant 1, p_1 < 1, p_0 + p_1 = 1$, 则 $f(0) = p_0$,

$$\begin{aligned} f_1(0) &= f(f(0)) = f(p_0) = p_0 + p_0 p_1, \cdots, \\ f_{n+1}(0) &= f(f_n(0)) = p_0 + p_0 p_1 + \cdots + p_0 p_1^n \\ &= p_0 \left( \frac{1 - p_i^{n+1}}{1 - p_1} \right) \quad (n \geqslant 0). \end{aligned}$$

所以 $\rho = \lim\limits_{n \to \infty} f_n(0) = \dfrac{p_0}{1 - p_1} = 1$.

(3) 设 $\infty \geqslant \mu > 1$, 则必有 $p_0 + p_1 < 1$. 所以 $f'(s)$ 在 $[0,1)$ 内严格上升. 如果我们能证: 存在 $b < 1$, 使

$$f(s) < s \quad (\forall s \in (b, 1)), \tag{4.17}$$

则

$$f(f_n(s)) < s, \quad \forall s \in (b, 1).$$

令 $s \downarrow b$ 得

$$b \geqslant f(f_n(b)) = f_{n+1}(b).$$

所以

$$\rho = \lim_{n \to \infty} f_n(0) \leqslant \lim_{n \to \infty} \sup f_n(b) \leqslant b < 1.$$

而 (1) 中已证 $\rho = f(\rho)$, 故 $\rho$ 是方程式 (4.14) 的一个解.

再证方程式 (4.14) 的解惟一. 若 $t_1, t_2$ 是方程式 (4.14) 的任意两个解. 不妨令 $0 \leqslant t_1 < t_2 < 1$, 则必存在 $u_1 \in (t_1, t_2), u_2 \in (t_2, 1)$, 使得

$$\begin{aligned} f'(u_1) &= \frac{f(t_2) - f(f_1)}{t_2 - t_1} = \frac{t_2 - t_1}{t_2 - t_1} = 1, \\ f'(u_2) &= \frac{f(1) - f(t_2)}{1 - t_2} = \frac{1 - t_2}{1 - t_2} = 1. \end{aligned}$$

所以

$$f'(u_1) = f'(u_2). \tag{4.18}$$

而今 $u_1 < u_2, f'(s)$ 在 $[0,1)$ 内严格上升, $u_i \in [0,1)$, 所以 (4.18) 不可能成立. 这就证明了方程式 (4.14) 的解是惟一的.

最后我们补充证明 (4.17). 因为

$$\lim_{s \uparrow 1} f'(s) = \mu > 1,$$

所以存在 $b<1$, 使 $f'(b)>1$. 但是当 $s\in(b,1)$ 时, 必有 $c$ 使 $1>c>s>b$, 且

$$\frac{f(1)-f(s)}{1-s}=f'(c)>f'(b)>1,$$

此即 $f(s)<s$. 定理证毕.

注意: 定理 4.2 的结论与直观意义是完全吻合的. 因为 $\mu=f'(1)=\sum\limits_{n=1}^{\infty}np_n$ 是一个质点在下一时刻所分裂成的质点数的期望. 当 $\mu\leqslant 1$ 且 $p_1<1$ 时, 该系统的质点总数的期望随时间的推移而不会增加, 且一质点经过一个单位时间后不可能总是保持一个. 在这种情况下, 该系统的质点总数 "迟早" 是会趋于 0 的. 当 $\mu>1$ 时, 即一个质点经过一个单位时间后, 其分裂出的质点数的期望大于 1, 当然该系统的质点总数不会以概率 1 而趋于 0.

定理 4.2 中假定了初始状态 $X_0\equiv 1$, 对于一般的 $X_0\equiv n_0$, 亦有类似的定理, 如下:

**定理 4.3** 设 $\{X_n:n\geqslant 0\}$ 是分枝链,$X_0\equiv n_0, f(s)$ 是其本原母函数, $\mu=f'(1)$. 令 $\overline{\rho}=\lim\limits_{n\to\infty}P(X_n=0)$ 是其绝灭概率. 再令

$$f_1(s)=f(s),\quad f_{n+1}(s)=f(f_n(s))=f_n(f(s))(n\geqslant 1),$$

$$\rho=\lim_{n\to\infty}f_n(0)\quad (\text{由定理 4.2 有 } \rho=f(\rho)),$$

则 $\overline{\rho}=\rho^{n_0}$.

(注意 $f_n(s)$ 不是 $X_n$ 的母函数, 而 $\rho$ 可根据定理 4.2 算出.)

**证** 令 $\overline{f}_n(s)$ 为 $X_n$ 的母函数, 则

$$\begin{aligned}\overline{f}_1(s)&=f(s)^{n_0},\\ \overline{f}_{n+1}(s)&=\overline{f}_n(f(s))=\overline{f}_{n-1}(f(f(s)))\\ &=\cdots=\overline{f}_1(f_n(s)).\end{aligned}$$

所以

$$\begin{aligned}\overline{\rho}&=\lim_{n\to\infty}\overline{f}_n(0)=\lim_{n\to\infty}\overline{f}_1(f_{n-1}(0))\\ &=\overline{f}_1\left(\lim_{n\to\infty}f_{n-1}(0)\right)=\overline{f}_1(\rho)\\ &=f(\rho)^{n_0}=\rho^{n_0}.\end{aligned}$$

**定理 4.4** 设 $\{X_n:n\geqslant 0\}$ 是分枝链, $X_0\equiv n_0, f(s)=\sum\limits_{n=0}^{\infty}p_ns^n$ 是其本原母函数, $\mu=f'(1)\geqslant 1, p_0+p_1<1, \sigma^2=f''(1)+f'(1)-(f'(1))^2>0$. 令

$$M_n=\boldsymbol{E}(X_n|X_n>0),\quad \zeta_n=\frac{X_n}{\boldsymbol{E}(X_n)},\quad \eta_n=\frac{X_n}{M_n},$$

则

$$(1)\ \lim_{n\to\infty} P(\eta_n \leqslant x|\eta_n>0)=\begin{cases} G_1(x), & 当\ \mu=1, f''(1)\neq 0, f'''(1)\ 有限, \\ G(x), & 当\ \mu>1, \sigma^2>0, \end{cases}$$

其中 $G_1(x)$ 是强度为 1 的负指数分布, $G(x)$ 是连续型分布函数其特征函数 $\varphi(t)$ 满足

$$\begin{cases} f((1-\rho)\varphi(t)+\rho)=(1-\rho)\varphi(\mu t)+\rho, \\ \varphi'(0)=i, \end{cases} \tag{4.19}$$

其中 $\rho$ 如定理 4.3 中所定义;

$$(2)\ \lim_{n\to\infty}\zeta_n=\zeta,\quad [\mathrm{P.}]\quad (当 \mu>1), \tag{4.20}$$

其中 $\zeta$ 的分布函数 $H(x)$ 在 0 有跃度 $\rho$, 在其他地方均连续;

$$(3)\ \lim_{n\to\infty} P\left(\frac{X_n-\mu X_{n-1}}{\sigma\sqrt{X_{n-1}}}<x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{1}{2}t^2}dt \quad (x\in\mathbf{R}). \tag{4.21}$$

**证** 参见 [41] p.20~21. 我们这里的 $\mu,\rho,X_n$ 分别是那里的 $a,\lambda,\xi_n$.

这条定理部分地解决了分枝链的极限分布的问题, 而定理 4.3 解决了分枝链的绝灭问题. 定理 4.1 解决了分枝链的爆炸问题, 因此在定理 4.1(3) 中有 $\mu_n \stackrel{\text{def.}}{=\!=} \boldsymbol{E}(X_n)=n_0\mu^n\to\infty$ (当 $\mu>1$ 时).

本节到此为止, 都是讨论离散时间的分枝链. 其实在实际问中, 所谓 "一个单位时间", 只不过是为了简化问题而说的. 在大数场合, 时间必须考虑是连续的. 因此, 我们有必要引进连续时的分枝过程. 在 (4.4) 中, 我们已经看到了分枝链作为一类特殊 Markov 链的概率特征. 因此, 我们就把 (4.4) "连续化" 作为分枝过程的定义.

**定义 4.3** 称时齐的可数状态的 Markov 过程 $X=\{X(t):t\in[0,\infty)\}$ 是**分枝过程**, 如果 $X$ 的转移矩阵 $P(t)=(p_{i,j}(t),i,j\geqslant 0)$ 满足:

$$\begin{cases} p_{i,j}(t)=\displaystyle\sum_{\sum_{s=1}^{i} j_s=j}\prod_{s=1}^{i}p_{1,j_s}(t) & (t\geqslant 0, i>0, j\geqslant 0), \\ p_{0,j}=\delta_{0,j} & (\delta_{i,j}\ 之定义如前). \end{cases} \tag{4.22}$$

撇开概率空间及随机过程, 纯分析地说, 任意一个满足 (4.22) 的 (准) 转移矩阵 $P(t)$ 都称为分枝 (准) 转移矩阵.

**命题 4.1** 设 $P(t)$ 是任何一个标准的分枝准转移矩阵, $P'(0)=Q=(q_{i,j},i,j\geqslant 0)$, 则

$$q_{i,j}=\begin{cases} 0, & 当\ j<i-1, \\ iq_{1,j-i+1}, & 当\ j\geqslant i-1, \end{cases} \tag{4.23}$$

此处 $q_i=-q_{i,i}$ 可以为 $\infty$.

**证** 由 (4.22) 式立即可推出 (4.23) 成立.

**定义 4.4** 设 $Q = (q_{i,j}, i, j \geqslant 0)$ 是一个转移强度矩阵 (自然要求 $q_i = -q_{i,i} < \infty, \forall i \geqslant 0$), 如果它满足 (2.23), 则称 $Q$ 是一个**分枝转移强度矩阵**.

由命题 4.1 看出: 对任何分枝准转移矩阵 $P(t)$ 来说, 只要 $p'_{i,i}(0) > -\infty(\forall i \geqslant 0)$, 则 $Q \stackrel{\text{def.}}{=\!=} P'(0)$ 是一个分枝转移强度矩阵.

**定理 4.5** 任给一个分枝转移强度矩阵 $Q = (q_{i,j}, i, j \geqslant 0)$, 令 $f(s) = \sum\limits_{j=0}^{\infty} q_{1,j} s^j$, 则

(1) 任何一个分枝 $Q$ 过程 $P(t) = (p_{i,j}(t), i, j \geqslant 0)$ 的母函数 $F(t,x) = \sum\limits_{j=0}^{\infty} p_{1,j}(t) x^j (|x| \leqslant 1)$ 必满足

$$\begin{cases} \dfrac{\partial F(t,x)}{\partial t} = f(F(t,x)), \\ F(0,x) = x \end{cases} \quad (|x| \leqslant 1, t \geqslant 0) \tag{4.24}$$

(从而最多只有一个分枝 $Q$ 过程);

(2) (4.24) 的解 $F(t,x)$ 是分枝 $Q$ 过程 $P(t)$ 的母函数的充分必要条件是

$$\left(\frac{\partial^k}{\partial x^k} F(t,x)\right)_{x=0} \geqslant 0, \quad F(t,1) \leqslant 1 \quad (t \geqslant 0, k = 0, 1, 2, \cdots); \tag{4.25}$$

(3) 对任一分枝转移强度矩阵 $Q$, 存在唯一一个分枝 $Q$ 过程, 它就是最小 $Q$ 过程.

证明参见 [38] 第二编定理 6.3.

作为这一节的结束, 我们再介绍一个连续时间的分枝过程 $X = \{X(t) : t \in [0, \infty)\}$ 的极限定理, 它类似于定理 2.2 中离散时间的分枝链的情形, 只不过当时的条件是加在分枝链 $\{X_n : n = 0, 1, 2, \cdots\}$ 的本原母函数 $f(s) = \sum\limits_{m=0}^{\infty} p_m s^m$ 上, 而今对连续时的分枝过程 $X = \{X(t) : t \in [0, \infty)\}$, 条件是加在分枝转移强度矩阵 $Q$ 上而已.

**定理 4.6** 设 $Q = (q_{i,j}, i, j \geqslant 0)$ 是一个分枝转移强度矩阵, $f(s) = \sum\limits_{j=0}^{\infty} q_{1,j} s^j$ 为其母函数, $P(t)$ 是一个分枝 $Q$ 过程.

(1) 若 $f(s) \equiv 0 (\forall |s| \leqslant 1)$, 则 $P(t) \equiv I$ 是单位矩阵, 更有 $\lim\limits_{t \to \infty} P(t) = I$;

(2) 若 $f(s) \neq 0(|s| \leqslant 1)$, 从而 $q_{1,1} < 0$, 则

$$\Pi = \lim_{t \to \infty} P(t) = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ \rho & 0 & 0 & \cdots \\ \rho^2 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix},$$

其中 $\rho$ 是 $f(x) = 0$ 的最靠近 0 的那一个非负根, 而且 $f'(1) > 0$ 时 $\rho < 1$. 特别地, 当 $f(1) = 0$ 时, 即 $Q$ 保守, 则 $\rho = 1$ 的充分必要条件是

$$f'(1) \stackrel{\text{def.}}{=\!=} \sum_{j=1}^{\infty} j q_{1,j} \leqslant 0.$$

证明请参见 [38] 第二编定理 6.4.

## §5 生灭过程与随机服务

本节主要讲 Markov 过程在排队论中的应用及生灭过程的背景. 生灭过程的一些证明很长的定理, 略去了证明, 但标明出处.

在这一节中, 我们主要讲述并研究随机服务系统中的一些基本概念及其有关问题, 以此作为背景, 再引进一类应用面很广的随机过程 —— 生灭过程. 当然, 在随机服务系统中, 应用到的随机过程理论, 不仅是生灭过程、更新过程, 一般 Markov 过程也常用到. 另一方面,生灭过程的应用, 也不只是用在随机服务系统中. 随机服务系统中的理论问题, 最早是从 "排队论" 中提出来的. 经过后来的发展, 远远超越了 "排队论" 的范畴. 但是为了尊重历史, 也是为了语言通俗, 我们在这一节中经常用 "排队论" 来代表随机服务系统中的理论问题. 一些名词、术语也借用排队论中的名词与术语, 其实它们都可以抽象为更广泛的问题.

排队论主要研究各类排队现象以便合理解决 "顾客" 与 "服务机构" 的供求矛盾, 这里顾客与服务机构都是抽象的, 视具体问题的不同而有不同的含义. 例如人们到售票站去购票, "人" 就是顾客, "售票窗口" 就是服务机构; 再如飞机要求在机场降落, "飞机" 就是顾客, 而 "跑道" 就是服务机构; 又如人们打电话, "电话呼叫" 就是顾客, 而 "电话交换台" 就是服务机构. 尽管顾客、服务机构、排队现象形形色色, 但从理论上来概括, 要构成一个完整的 "服务系统", 都必须具备下列三个组成部分:

(1) 输入过程. 是指来到服务机构请求为之服务的顾客到来的规律.

(2) 服务规则. 是指各种类型的 "顾客" 来到服务机构时, 按什么原则什么次序接受服务.

(3) 服务机构. 是指服务机构内部的设备、能力、服务速度等等.

下面我们逐个来简单介绍一下这三个组成部分.

(1) 输入过程

一般而言, 有下列几种输入. 令 $X_t$ 是 $[0,t)$ 内来到服务机构请求服务的顾客数, $\tau_n$ 是第 $n$ 个顾客到来的时刻, $\tau_1=T_1,\cdots,\tau_n=\sum\limits_{k=1}^{n}T_k(n\geqslant 1)$.

(a) 定长输入. 即是存在一个常数 $\alpha$, 使 $P(T_n=\alpha)=1(n\geqslant 1)$, 即相邻到来的两个顾客的间隔恒为 $\alpha$, 亦即

$$A(t)\equiv P(T_n\leqslant t)=\begin{cases}0, & \text{当 } t\leqslant\alpha,\\ 1, & \text{当 } t>\alpha.\end{cases}\tag{5.1}$$

(b) Poisson 输入. 即是 $\{X_t:t\in[0,\infty)\}$ 是一个 Poisson 过程 (强度为 $\lambda$). 可以证明 (参见定理 5.1): $\{X_t\}$ 是强度为 $\lambda$ 的 Poisson 过程的充分必要条件是: $\{X_t<n\}=\{\tau_n\geqslant t\}$, 其中 $\tau_n=\sum\limits_{k=1}^{n}T_k,\{T_n\}$ 是一串相互独立且具有共同的负指数分布:

$$P(T_n\leqslant t)=\begin{cases}1-e^{-\lambda t}, & \text{当 } t>0,\\ 0, & \text{当 } t\leqslant 0.\end{cases}\tag{5.2}$$

(c) Erlang 输入. 即是 $\{T_n:n=0,1,\cdots\}$ 是一串相互独立、具有公共分布函数的随机变量, 其密度函数为

$$a(t)=\begin{cases}0, & \text{当 } t\leqslant 0,\\ \dfrac{\lambda(\lambda t)^{k-1}}{(k-1)!}e^{-t\lambda}, & \text{当 } t>0\end{cases}\tag{5.3}$$

($\lambda>0,k$ 是一个正整数), 亦即 $T_n$ 服从参数为 $(k,\lambda)$ 的 $\Gamma$ 分布.

注意: 当 $k=1$, Erlang 输入即为 Poisson 输入.

(d) 一般输入. 即是 $\{T_n:n=0,1,2,\cdots\}$ 是一串相互独立、相同分布的随机变量.

(a),(b),(c) 都是 (d) 的特例.

(2) 服务规则

一般而言, 有三大类.

(a) 损失制. 即是顾客到达时, 如果服务机构内的所有服务设施都被占, 该顾客就自动离去, 不再排队等候. 例如我国通用的电话系统就是采取损失制的服务规则. 一次电话呼叫 (来到一个顾客), 当所有的电话线被占用时, 只此呼叫就消失.

(b) 等待制. 顾客到达时, 若服务机构内的所有服务设施都被占, 此顾客就加入等待服务的排队行列. 这种服务规则中, 如果细分, 还有先到先服务、优先权服务、随机服务等等方式.

(c) 混合制. 即兼有消失制与等待制两种含义. 例如规定: 等待服务的顾客不超过 $N$ (确定的一个正整数) 时, 新来的顾客就加入排队行列, 若队长超过 $N$, 新来的顾客就离去. 再如: 顾客在队伍中等待的时间不超过 $T$, 若超过 $T$ 就离去.

(3) 服务机构. 主要是指服务设施的数目 (例如机场有 $N$ 条跑道, 电话交换台有 $M$ 条线路, 某火车站有 $k$ 个售票窗口, 等等) 及每一个服务设施的服务速度. 关于每一个服务设施的服务速度, 大致上又有下列几种 (令 $V$ 是每个顾客的服务时间):

(a) 定长分布. 即是 $P(V=\beta)=1$, 亦即

$$B(t) \equiv P(V<t)=\begin{cases}0, & \text{当 } t \leqslant \beta, \\ 1, & \text{当 } t>\beta.\end{cases} \tag{5.4}$$

(b) 负指数分布. 即各顾客的服务时间相互独立且具有相同的负指数分布:

$$B(t) \equiv P(V \leqslant t)=\begin{cases}1-e^{-\mu t}, & \text{当 } t \geqslant 0, \\ 0, & \text{当 } t<0,\end{cases} \tag{5.5}$$

其中 $\mu$ 是一正数. 平均服务时间为

$$\boldsymbol{E}(V)=\int_0^{\infty} x \mu e^{-\mu x} d x=\frac{1}{\mu}. \tag{5.6}$$

(c) Erlang 分布. 即各顾客的服务时间相互独立、具有相同的 Erlang 分布, 其密度函数为

$$b(t)=\begin{cases}\dfrac{k \mu(k \mu t)^{k-1}}{(k-1)!} e^{-\mu k t}, & \text{当 } t>0, \\ 0, & \text{当 } t \leqslant 0,\end{cases} \tag{5.7}$$

其中 $\mu>0, k$ 为正整数.

(d) 一般分布. 即各顾客的服务时间相互独立、具有相同的分布函数.

在服务系统中, 几种重要的数量指标为:

(1) 等待时间. 即从顾客到达服务机构至他开始被服务这段时间. 从顾客的立场出发, 当然等待时间愈短愈好. 但是等待时间愈短, 势必要求服务机构的服务设施多、速度快. 如何调整这一矛盾, 以使顾客与服务机构双方都满意, 这正是排队论要解决的矛盾之一.

(2) 队长. 即是在等待服务的顾客数. 这也是顾客与服务机构双方都关心的问题. 就服务机构而言, 正确估计出队长, 对设计等待空间 (如候机室、停车场、仓库、码头等等) 的大小、服务效率的快慢, 都有很大关系. 等待空间设计过大, 则造成设施闲置浪费, 等待空间过小, 则容纳不了等待服务的顾客. 服务效率太

低, 服务系统前经常出现长队, 则引起顾客厌烦而离去, 或下次不再来. 服务效率太高, 势必增加服务设施. (当然, 加强科学管理, 在不增加设施的情况下, 也可以提高服务效率, 但我们在这里假设科学管理已达到较理想的程度, 而从数学的角度去研究服务设施的数量与服务效率高低的关系.) 就顾客而言, 当然希望队伍不太长为好, 这样等候服务的时间会少一些.

(3) 忙期. 即服务机构连续繁忙的时间. 这直接关系到服务机构中的服务人员, 服务设施的工作强度. 也涉及顾客的利益, 例如呼叫一次电话, 总希望不至发生所有可使用的线路处于忙期. 因为所有可使用的线路都处于忙期, 电话就打不通了.

下面我们首先研究输入过程.

**定理 5.1** 输入过程 $\{X_t : t \in [0,\infty)\}$ 是强度为 $\lambda$ 的 Poisson 过程的充分必要条件是: $\forall n \geqslant 1, t \in [0,\infty)$ 有: $\{X_t < n\} = \{\tau_n \geqslant t\}$, 其中 $\tau_n = \sum\limits_{k=1}^{n} T_k, \{T_n, n \geqslant 1\}$ 是相互独立服从共同分布的随机变量序列, 且其公共分布为以 $\lambda$ 为参数的负指数分布:

$$P(T_n \leqslant t) = \begin{cases} 1 - e^{-\lambda t}, & \text{当 } t > 0, \\ 0, & \text{当 } t \leqslant 0. \end{cases} \tag{5.8}$$

**证** 必要性. 设 $\{X_t : t \in [0,\infty)\}$ 是一个强度为 $\lambda$ 的 Poisson 输入过程, $\tau_n$ 是第 $n$ 个顾客到达的时刻, $T_n = \tau_n - \tau_{n-1} (n \geqslant 1, \tau_0 \equiv 0), \{X_t = n\} = \{\tau_n < t \leqslant \tau_{n+1}\}, \{X_t \geqslant n\} = \{\tau_n < t\} (t \in [0,\infty), n = 0, 1, 2, \cdots)$. 记 $X_t = X(t)$. 推证 $T_1, T_2, \cdots, T_k$ 相互独立, 且每个 $T_i$ 具有形如 (5.8) 的负指数分布. 为使符号简单起见, 只证 $k = 2$ 的场合.

任取 $t_1 > 0, t_2 > 0$, 把 $[0, t_1) n$ 等分, 令

$$A_n = \bigcup_{j=0}^{n-1} \left\{ X\left(\frac{(j+1)t_1}{n}\right) - X\left(\frac{jt_1}{n}\right) \leqslant 1 \right\},$$

$$B_n = \bigcap_{j=0}^{n-1} \left\{ X\left(\frac{(j+1)t_1}{n}\right) - X\left(\frac{jt_1}{n}\right) > 1 \right\},$$

则 $A_n$ 与 $B_n$ 互为补集, 且

$$A_n \bigcap \{X(t_1) \geqslant 1\} \subset \bigcup_{j=0}^{n-1} \left\{ X\left(\frac{jt_1}{n}\right) = 0, X\left(\frac{(j+1)t_1}{n}\right) = 1 \right\},$$

所以由 $\left\{X\left(\frac{(j+1)t_1}{n}\right)=1\right\}\subset\left\{\tau_1<\frac{(j+1)t_1}{n}\right\}$ 可得

$$\begin{aligned}
&A_n\bigcap\{X(t_1)\geqslant 1\}\bigcap\{T_2<t_2\}\\
\subset&\bigcup_{j=0}^{n-1}\left\{X\left(\frac{jt_1}{n}\right)=0,X\left(\frac{(j+1)t_1}{n}\right)=1,\tau_1<\frac{(j+1)t_1}{n},T_2<t_2\right\}\\
=&\bigcup_{j=0}^{n-1}\left\{X\left(\frac{jt_1}{n}\right)=0,X\left(\frac{(j+1)t_1}{n}\right)=1,X\left(t_2+\frac{(j+1)t_1}{n}\right)\geqslant 2\right\}\\
=&\bigcup_{j=0}^{n-1}\left\{X\left(\frac{jt_1}{n}\right)=0,X\left(\frac{(j+1)t_1}{n}\right)-X\left(\frac{jt_1}{n}\right)=1,\right.\\
&\left.X\left(t_2+\frac{(j+1)t_1}{n}\right)-X\left(\frac{(j+1)t_1}{n}\right)\geqslant 1\right\}
\end{aligned}\tag{5.9}$$

故由 $\{X(t):t\in[0,\infty)\}$ 是强度为 $\lambda$ 的 Poisson 过程可得

$$\begin{aligned}
&P(T_1<t_1,T_2<t_2)\\
=&P(X(t_1)\geqslant 1,T_2<t_2)\\
\leqslant&P(A_n,X(t_1)\geqslant 1,T_2<t_2)+P(B_n)\\
\leqslant&\sum_{j=0}^{n-1}P\left(X\left(\frac{jt_1}{n}\right)=0,X\left(\frac{(j+1)t_1}{n}\right)-X\left(\frac{jt_1}{n}\right)=1,\right.\\
&\left.X\left(\frac{(j+1)t_1}{n}+t_2\right)-X\left(\frac{(j+1)t_1}{n}\right)\geqslant 1\right)+P(B_n)\\
=&\sum_{j=0}^{n-1}e^{-\lambda jt_1/n}\cdot\frac{\lambda t_1}{n}e^{-\lambda t_1/n}\cdot(1-e^{-\lambda t_2})+P(B_n)\\
=&(1-e^{-\lambda t_1})(1-e^{-\lambda t_2})+o(1)\quad(n\to\infty).
\end{aligned}\tag{5.10}$$

在 (5.10) 中令 $n\to\infty$ 得

$$P(T_1<t_1,T_2<t_2)\leqslant(1-e^{-\lambda t_1})(1-e^{-\lambda t_2}).\tag{5.11}$$

下面证明 (5.11) 的反向不等式. 仿 (5.10) 有

$$P(T_1<t_1,T_2<t_2)\geqslant\sum_{j=0}^{n-1}e^{-\frac{\lambda jt_1}{n}}\cdot\frac{\lambda t_1}{n}e^{-\frac{\lambda t_1}{n}}\cdot(1-e^{-\lambda\left(t_2-\frac{t_1}{n}\right)}).$$

在上式中令 $n\to\infty$ 得

$$P(T_1<t_1,T_2<t_2)\leqslant(1-e^{-\lambda t_1})(1-e^{-\lambda t_2}).\tag{5.12}$$

由 (5.11), (5.12) 得

$$P(T_1 < t_1, T_2 < t_2) = (1 - e^{-\lambda t_1})(1 - e^{-\lambda t_2}) \quad (t_1 > 0, t_2 > 0). \tag{5.13}$$

在 (5.13) 中令 $t_2 \uparrow \infty$ 得

$$P(T_1 < t_1) = (1 - e^{-\lambda t_1}) \quad (t_1 > 0). \tag{5.14}$$

在 (5.13) 中令 $t_1 \uparrow \infty$ 得

$$P(T_2 < t_2) = (1 - e^{-\lambda t_2}) \quad (t_2 > 0). \tag{5.15}$$

由 (5.13), (5.14), (5.15) 得

$$\begin{aligned} P(T_1 < t_1, T_2 < t_2) &= P(T_1 < t_1)P(T_2 < t_2) \\ &= (1 - e^{-\lambda t_1})(1 - e^{-\lambda t_2}) \quad (t_1 > 0, t_2 > 0). \end{aligned} \tag{5.16}$$

而当 $t_1, t_2$ 中有一个为 0 时, (5.16) 显然成立. 而当 $t \leqslant 0$ 时, 总有 $P(T_n < t) = 0$. 总之, $T_1$ 与 $T_2$ 相互独立, 且 $T_n$ 服从强度为 $\lambda$ 的负指数分布. 必要性得证.

充分性. 设 $\{T_n : n \geqslant 1\}$ 独立同分布且其公共分布强度为 $\lambda$ 的负指数分布. 则 $\tau_n = \sum\limits_{i=1}^{n} T_i$ 服从参数为 $(n, \lambda)$ 的 $\Gamma$ 分布, 即 $\tau_n$ 的密度函数为

$$a_n(t) = \begin{cases} \dfrac{\lambda(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda t}, & 当\ t > 0, \\ 0, & 当\ t \leqslant 0. \end{cases} \tag{5.17}$$

推证 $\{X(t) : t \in [0, \infty)\}$ 是一个强度为 $\lambda$ 的 Poisson 过程. 为此, 只需证明: 对任何正整数, 任何非负实数 $0 = t_0 < t_1 < \cdots < t_n$ 及任何非负整数 $i_1, \cdots, i_n$, 总有下列等式:

$$P\left(\bigcap_{k=1}^{n} \{X(t_k) - X(t_{k-1}) = i_k\}\right) = \prod_{k=1}^{n} e^{-\lambda(t_k - t_{k-1})} \cdot \frac{(\lambda(t_k - t_{k-1}))^{i_k}}{i_k!}. \tag{5.18}$$

因为由 (5.18) 及 $\{X(t) : t \in [0, \infty)\}$ 的定义可容易验证它是强度为 $\lambda$ 的 Poisson 过程. 至于 (5.18) 的证明可以对 $n$ 作归纳法并区分 $i_1, \cdots, i_n$ 的各种取值情况来证明.

**定义 5.1** 设 $X = \{X_t : t \in [0, \infty)\}$ 是一个以 $E = \{0, 1, 2, \cdots\}$ 为状态空间以 $P(t) = (p_{i,j}(t), i, j \in E)$ 为转移矩阵的时齐的 Markov 过程. $Q = (q_{i,j}, i, j \in E) = P'(0)$ 是其转移强度矩阵, 如果

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}, \tag{5.19}$$

其中 $\lambda_0 \geqslant 0; \lambda_1, \lambda_2, \cdots > 0; \mu_1, \mu_2, \cdots > 0$ (即 $q_{0,0} = -\lambda_0, q_{0,1} = \lambda_0, q_{0,j} = 0$ (当 $j \geqslant 2$). 而当 $i \geqslant 1$ 时, $q_{i,i} = -(\lambda_i + \mu_i), q_{i,i+1} = \lambda_i, q_{i,i-1} = \mu_i, q_{i,j} = 0$ (当 $|i-j| > 1$)), 则称 $X$ 是一个**生灭过程**.

撇开概率背景, 纯分析地说, 任意一个形如 (5.19) 的矩阵都称为一个**生灭转移强度矩阵**. 显然, 它是保守的.

下面我们结合随机服务系统来看一个生灭过程的例子.

**例 5.1** 设有一轮船码头, 在 $[0,t)$ 时间区间内来到码头要求装卸货物的轮船数为 $X_t, \{X_t : t \in [0,\infty)\}$ 构成一个以 $\lambda$ 为强度的 Poisson 过程, 再设每条轮船装卸货物的时间 $T$ 服从参数为 $\mu$ 的负指数分布, 而且各条轮船装卸货物的时间相互独立, 且与到来的轮船亦相互独立. 令 $\eta_t$ 是时刻 $t$ 码头内之轮船数. 则 $\{\eta_t : t \geqslant 0\}$ 是一个生灭过程, 其转移密度矩阵 $P'(0)$ 是一个形如 (5.19) 的生灭转移强度矩阵.

**证** 令 $A_k(t) = \{$在 $[0,t)$ 内共有 $k$ 条轮船离开码头$\}, B_k(t) = \{X_t = k\} = \{$在 $[0,t)$ 内共有 $k$ 条轮船到达码头$\}$ $(k = 0,1,2,\cdots)$. $T_i$ 是第 $i$ 条轮船装卸货物所需的时间 $(i = 1,2,\cdots)$. 由于 $T$ 服从负指数分布, 所以对任一条轮船, 若已知 "它在时刻 $t_1$ 到达码头, 而且时刻 $t_2$ 尚未离开码头 $(t_2 \geqslant t_1)$", 而且在时刻 $t_2 + t$ 仍未离开码头的概率

$$P(T > t_2 - t_1 + t | T > t_2 - t_1) = \frac{P(T > t_2 - t_1 + t)}{P(T > t_2 - t_1)} = e^{-\mu t} = P(T > t)$$

不依赖于 $t_1$ 和 $t_2$. 因此, 时刻 0 在码头的任何一条轮船 (这里把观察时刻当作时刻起点 0), 在时刻 $t$ 尚未离开码头的概率为

$$P(T > t) = e^{-\mu t}.$$

所以用独立性假设可知

$$P(A_0(t)B_0(t)|\eta_0 = i) = e^{-\lambda t} \bullet e^{-i\mu t}. \tag{5.20}$$

显然, 由 $T$ 服从负指数分布知

$$\lim_{t\to 0+} P\left(\bigcup_{r=1}^{i} \{T_r < t\}\right) = 0 \quad (\forall i \geqslant 1). \tag{5.21}$$

而

$$\begin{aligned} p_{i,i}(t) &\overset{\text{def.}}{=\!=} P(\eta_t = i|\eta_0 = i) = \frac{P\left(\{\eta_0 = i\} \bigcap \left(\bigcup\limits_{k=0}^{\infty} B_k(t) \bigcap A_k(t)\right)\right)}{P(\eta_0 = i)} \\ &= \sum_{k=0}^{\infty} P(B_k(t)A_k(t)|\eta_0 = i), \end{aligned} \tag{5.22}$$

但是, 由 (5.20) 知

$$\lim_{t\to 0+}\frac{P(B_0(t)A_0(t)|\eta_0=i)-1}{t}=-(\lambda+i\mu). \tag{5.23}$$

若令 $v_k(t)=P(X_t=k)$, 则

$$\begin{aligned}0&\leqslant\frac{\sum\limits_{k=1}^{\infty}P(B_k(t)A_k(t)|\eta_0=i)}{t}\\&\leqslant\frac{P(B_1(t)A_1(t)|\eta_0=i)}{t}+\frac{\sum\limits_{k=2}^{\infty}v_k(t)}{t}\\&\leqslant\frac{P\left(X_t=1,\bigcup\limits_{r=1}^{i+1}\{T_r<t\}|\eta_0=i\right)}{t}+o(1)\\&=\lambda e^{-\lambda t}P\left(\bigcup_{r=1}^{i+1}\{T_r<t\}\right)+o(1),\end{aligned} \tag{5.24}$$

由 (5.21)~(5.24) 得

$$\lim_{t\to 0+}\frac{p_{i,i}(t)-1}{t}=-(\lambda+i\mu)\quad (i\in E). \tag{5.25}$$

仿 (5.22) 有

$$p_{i,i+1}(t)=\sum_{k=1}^{\infty}P(B_k(t)A_{k-1}(t)|\eta_0=i)\quad (i\in E). \tag{5.26}$$

仿 (5.25) 之证明方法有

$$\lim_{t\to 0+}\frac{p_{i,i+1}(t)}{t}=\lambda\quad (i\in E). \tag{5.27}$$

当 $i=1,2,\cdots,j=i-1$ 时, 仿 (5.22) 有

$$p_{i,i-1}(t)=\sum_{k=1}^{\infty}P(B_{k-1}(t)A_k(t)|\eta_0=i). \tag{5.28}$$

仿 (5.25) 之证明方法, 由 (5.28) 有

$$\lim_{t\to 0+}\frac{p_{i,i-1}(t)}{t}=i\mu\quad (i=1,2,\cdots). \tag{5.29}$$

而对任何 $|i-j|>1$, 仿 (5.25) 之证明方法有

$$\lim_{t\to 0+}\frac{p_{i,j}(t)}{t}=0\quad (\forall |i-j|>1). \tag{5.30}$$

由 (5.25), (5.27), (5.29), (5.30) 得

$$P'(0) = Q$$
$$= \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \cdots \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

显然, $\{\eta_t : t \in [0,\infty)\}$ 是以 $P(t)$ 为转移矩阵、以 $Q$ 为转移强度矩阵、以 $E$ 为状态空间的时齐的 Markov 过程.

**定义 5.2** 若 $Q$ 是形如 (5.19) 的生灭转移强度矩阵, 当 $\mu_i \equiv 0(\forall i \geqslant 1)$ 时, 则称 $Q$ 是**纯生的**; 当 $\lambda_i \equiv 0(\forall i \geqslant 0)$ 时, 则称 $Q$ 是**纯灭的**. 如果 Markov 过程的转移强度矩阵是纯生 (纯灭) 转移强度矩阵时, 则称此 Markov 过程是**纯生 (纯灭) 过程.**

**定理 5.2** 设 $Q$ 是一个形如 (5.19) 的生灭转移强度矩阵, 则下列陈述等价:

(1) $\overline{P}(t)$ 是标准的转移矩阵 (此处 $\overline{P}(t)$ 是本章定理 2.4 中所定义的最小的 $Q$ 过程), 从而 $Q$ 过程是唯一的;

(2)《$(\lambda I - Q)y = 0, y \geqslant 0, \sup\limits_{i\geqslant 0} y_i < \infty$》只有零解 ($\lambda$ 为任一大于 0 的实数);

(3)
$$\sum_{n=1}^{\infty}\sum_{k=1}^{n} \frac{\mu_{k+1}\mu_{k+2}\cdots\mu_n}{\lambda_k\lambda_{k+1}\lambda_{k+2}\cdots\lambda_n} = \infty. \tag{5.31}$$

证明参见 [38] 第二编定理 5.1.

这个定理说明: 对生灭过程来说, 只要其转移强度矩阵 $Q$ 满足 (5.31), 其 $Q$ 过程是唯一的, 所以此生灭过程的概率规律性完全由 $Q$ 所决定.

**推论 5.1** 若 $Q = (q_{i,j}, i, j \geqslant 0)$ 是一个纯生转强阵, 即 $q_{i,i} = -\lambda_i, q_{i,i+1} = \lambda_i, q_{i,j} = 0$ (当 $j \neq i, j \neq i+1), \lambda_0 \geqslant 0, \lambda_i > 0(i \geqslant 1)$, 则下列陈述等价:

(1) 最小 $Q$ 过程 $\overline{P}(t)$ 是标准的转移矩阵;

(2)《$(\lambda I - Q)y = 0, y \geqslant 0, \sup\limits_{i\geqslant 0} y_i < \infty$》只有零解 (其中 $\lambda$ 为任一正实数);

(3)
$$\sum_{i=1}^{\infty} \frac{1}{\lambda_i} = \infty. \tag{5.32}$$

**推论 5.2** 设 $X = \{X_t : t \in [0,\infty)\}$ 是满足推论 5.1 中条件 (5.32) 的纯生

过程, $P(t)=(p_{i,j}(t),i,j\geqslant 0)$ 是其转移矩阵, 则

$$p_{i,j}(t)=\begin{cases}0, & \text{当 } j<i\\ e^{-\lambda_i t}, & \text{当 } j=i,\\ e^{-\lambda_j t}\int_0^t e^{\lambda_j s}\lambda_{j-1}p_{i,j-1}(s)ds, & \text{当 } j>i.\end{cases}$$

**推论 5.3** 设 $X$ 是推论 5.2 中的纯生过程. 若 $\lambda_i\equiv\lambda>0(\forall i\geqslant 0)$ 则 $X$ 必为强度为 $\lambda$ 的 Poisson 过程, 反之亦真.

**定理 5.3** 设 $X=\{X_t:t\in[0,\infty)\}$ 是一个生灭过程, 其转移强度矩阵 $Q=(q_{i,j},i,j\geqslant 0)$ 由 (5.19) 所给出. $\overline{P}(t)=(\overline{p}_{i,j}(t),i,j\geqslant 0)$ 是本章定理 2.4 中所定义的最小 $Q$ 过程, $P(t)=(p_{i,j}(t),i,j\geqslant 0)$ 是 $X$ 的转移矩阵, ($P(t)$ 和 $\overline{P}(t)$ 并未给出, 而 $Q$ 却是已知的) 令

$$\pi_{i,j}=\lim_{t\to\infty}p_{i,j}(t)\quad(i,j\geqslant 0),$$
$$\overline{\pi}_{i,j}=\lim_{t\to\infty}\overline{p}_{i,j}(t)\quad(i,j\geqslant 0).$$

(甲) 设 $\lambda_0>0$, 令

$$\rho_0=1,\quad \rho_n=\frac{\lambda_0\lambda_1\cdots\lambda_{n-1}}{\mu_1\mu_2\cdots\mu_n}\quad(n\geqslant 1).$$

(1) 若 $\sum\limits_{n=0}^{\infty}\rho_n=\infty$, 则 $\overline{\pi}_{i,j}=0(\forall i,j\geqslant 0)$.

(2) 若 $\sum\limits_{n=0}^{\infty}\rho_n<\infty$, $\sum\limits_{n=0}^{\infty}\dfrac{1}{\lambda_n\rho_n}=\infty$, 则

$$\overline{\pi}_{i,j}=\frac{\rho_j}{\sum\limits_{n=0}^{\infty}\rho_n}\quad(i\geqslant 0,j\geqslant 0).$$

(3) 若 $\sum\limits_{n=0}^{\infty}\rho_n<\infty$, $\sum\limits_{n=0}^{\infty}\dfrac{1}{\lambda_n\rho_n}<\infty$, 则

$$\overline{\pi}_{i,j}=0\quad(i\geqslant 0,j\geqslant 0).$$

(乙) 设 $\lambda_0=0$, 令

$$\gamma=1+\frac{\mu_1}{\lambda_1}+\frac{\mu_1\mu_2}{\lambda_1\lambda_2}+\cdots,$$
$$y_0^{(0)}=1,\quad y_n^{(0)}=\frac{1}{\gamma}\sum_{k=n}^{\infty}\frac{\mu_1\cdots\mu_k}{\lambda_1\cdots\lambda_k}\quad(n\geqslant 1).$$

(1) 若 $\gamma=\infty$, 则

$$\overline{\pi}_{i,0}=1,\quad \overline{\pi}_{i,j}=0\quad(i\geqslant 0,j\geqslant 1).$$

(2) 若 $\gamma < \infty$, 则

$$\overline{\pi}_{i,0} = y_i^{(0)}, \quad \overline{\pi}_{i,j} = 0 \quad (i \geqslant 0, j \geqslant 1).$$

特别地, 若 (5.31) 成立, 则 $Q$ 过程是唯一的, 从而 $P(t) = \overline{P}(t)$, 所以

$$\pi_{i,j} = \lim_{t\to\infty} p_{i,j}(t) = \lim_{t\to\infty} \overline{p}_{i,j}(t) = \overline{\pi}_{i,j} \quad (i, j \geqslant 0).$$

于是 $\pi_{i,j}$ 亦由定理 5.3 中诸款而计算出来了.

证明参见 [38] 第二编定理 5.7 与定理 5.8.

定理 5.3 告诉我们: 只要知道生灭过程的转移强度矩阵 $Q$ (不必知道中间对象 $P(t) = (p_{i,j}(t), i, j \geqslant 0)$), 就可以直接算出

$$\pi_{i,j} = \lim_{t\to\infty} p_{i,j}(t).$$

这在实际问题中是非常有用的.

我们这里只把生灭过程作为随机服务系统的理论基础之一而简单地予以介绍, 其实生灭过程的内容非常丰富. 有兴趣的读者可参看专著 [107] 或 [38].

现在我们来综合研究服务系统. 如前所述, 一个完整的服务系统包括三个组成部分:

(1) 输入过程; (2) 服务规则; (3) 服务机构.

对输入过程, 我们用 $M, D, E_k, GI$ 分别表示 Poisson 输入、定长输入、参数为 $(k, \lambda)$ 的 Erlang 输入 (即参数为 $(k, \lambda)$ 的 $\Gamma$ 输入) 和一般输入. 对 "服务规则", 都是等待制或损失制, 且先到先服务, 对于服务机构, 主要有两个指标, 一是服务设施的数量, 二是每台服务设施的服务速度. 服务设施的数量用 $n$ (有限个) 或 $\infty$ (无穷多个) 表示. 服务速度用 $G$ 表示各顾客的服务时间是一串相互独立具有公共分布的随机变量, 用 $M$ 表示各顾客的服务时间是一串相互独立的具有公共的负指数分布的随机变量, $D$ 表示各顾客的服务时间为定长, $E_k$ 表示各顾客的服务时间是一串相互独立的具有以 $(k, \lambda)$ 为参数的 Erlang (或称 $\Gamma$) 分布的随机变量. 而且我们总设各顾客的服务时间与输入过程相互独立. 这样一来, 对于一个服务系统, 只要用 4 个指标就可以刻画了, 例如 $M/M/n$ (等待) 表示服务系统的组成如下: 输入是 Poisson 输入, 服务规则是等待制且先到先服务, 服务设施数目是 $n$, 每台设施的服务速度 (即每个顾客所需的服务时间) 服从负指数分布. 再如 $GI/M/n$ (损失) 表示一般输入, 服务规则是先到先服务从而是损失制, 服务设施数为 $n$. 每台设施的服务速度服从负指数分布.

(一) $M/M/n$ (等待) 系统

在这一段中, 恒设

(1) 输入过程是以 $\lambda$ 为强度的 Poisson 过程;

(2) 服务规则是等待制先到先服务, 即是顾客到达时, 若有服务设施空闲, 则该顾客立即被接受服务; 若所有的 ($n$ 台) 服务设施均被占用, 则该顾客加入等待队伍, 依次等待接受服务;

(3) 服务机构中服务设施的数目为 $n$, 每个顾客的服务时间服从参数为 $\mu$ 的负指数分布, 从而每个顾客的平均服务时间为 $\frac{1}{\mu}$.

(4) 各顾客的服务时间是相互独立的, 且与输入过程也是相互独立的.

令 $\eta_t$ 表示时刻 $t$ 在该服务机构中正在服务的顾客数与等待服务的顾客数的总和. 则仿例 3.1 可证: $\{\eta_t : t \in [0,\infty)\}$ 是一个生灭过程, 其密度矩阵 $Q = (q_{i,j}, i,j \geqslant 0)$ 有如下形式:

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2). & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix},$$

其中 $\lambda_j \equiv \lambda, j = 0,1,2,\cdots$,

$$\mu_j = \begin{cases} j\mu, & \text{当 } j = 1,2,\cdots,n, \\ n\mu, & \text{当 } j = n+1, n+2, \cdots. \end{cases} \tag{5.33}$$

令 $P(t) = (p_{i,j}(t), i,j \geqslant 0)$ 是 $\{\eta_t : t \in [0,\infty)\}$ 的转移矩阵, $\overline{P}(t) = (\overline{p}_{i,j}(t), i,j \geqslant 0)$ 由本章定理 2.4 所定义.

**定理 5.4** 对于 $M/M/n$ (等待) 系统而言, 若 $\frac{\lambda}{n\mu} < 1$, 则 $\{\eta_t, t \in [0,\infty)\}$ 的转移矩阵 $P(t) = \overline{P}(t)$, 且

$$\lim_{t\to\infty} p_{i,j}(t) = \frac{\rho_j}{\sum\limits_{k=0}^{\infty} \rho_k} \quad (i \geqslant 0, j \geqslant 0), \tag{5.34}$$

其中

$$\rho_0 = 1, \rho_k = \frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k} \quad (k \geqslant 1), \tag{5.35}$$

$\lambda_j, \mu_j$ 如 (5.33) 所定义.

**证** 由于 $\dfrac{\lambda}{n\mu} < 1$, 所以

$$\begin{aligned}
&\sum_{m=1}^{\infty}\sum_{k=1}^{m}\frac{\mu_{k+1}\mu_{k+2}\cdots\mu_m}{\lambda_k\lambda_{k+1}\lambda_{k+2}\cdots\lambda_m}\\
\geqslant&\sum_{m=n+1}^{\infty}\sum_{k=n+1}^{m}\frac{\mu_{k+1}\mu_{k+2}\cdots\mu_m}{\lambda_k\lambda_{k+1}\lambda_{k+2}\cdots\lambda_m}\\
=&\frac{1}{\lambda}\sum_{m=n+1}^{\infty}\sum_{k=n+1}^{m}\left(\frac{n\mu}{\lambda}\right)^{m-k}=\infty.
\end{aligned}$$

因此由定理 5.2 知 $P(t) = \overline{P}(t)$. 又因为

$$\sum_{k=0}^{\infty}\rho_k=\sum_{k=0}^{n}\rho_k+\sum_{k=n+1}^{\infty}\left(\frac{\lambda}{\mu}\right)^n\frac{1}{n!}\left(\frac{\lambda}{n\mu}\right)^{k-n}<\infty,$$

$$\begin{aligned}
\sum_{k=0}^{\infty}\frac{1}{\lambda_k\rho_k}&=\sum_{k=1}^{\infty}\frac{\mu_1\mu_2\cdots\mu_k}{\lambda_0\lambda_1\lambda_2\cdots\lambda_k}+\frac{1}{\lambda_0}\\
&\geqslant\sum_{k=n+1}^{\infty}\frac{\mu_1\mu_2\cdots\mu_k}{\lambda_0\lambda_1\lambda_2\cdots\lambda_k}\\
&=\frac{1}{\lambda}\sum_{k=n+1}^{\infty}\left(\frac{\mu}{\lambda}\right)^n n!\left(\frac{n\mu}{\lambda}\right)^{k-n}\\
&=\infty,
\end{aligned}$$

所以, 由定理 5.3 (甲) 得知

$$\lim_{t\to\infty}p_{i,j}(t)=\frac{\rho_j}{\sum\limits_{k=0}^{\infty}\rho_k}\quad(i\geqslant 0, j\geqslant 0).$$

定理 5.4 证毕.

现在我们来分析一下定理 5.4.

(1) 条件 $\dfrac{\lambda}{n\mu} < 1$ 的意义. 我们知道 $\lambda$ 是输入过程 —— Poisson 过程的强度, 它的直观意义是 "单位时间内来到的顾客的平均数", 而 $\mu$ 是服务时间所遵从的负指数分布的参数, 它的直观意义是: $\dfrac{1}{\mu}$ 是每个顾客的 "平均服务时间", 从而 $\mu$ 是一台服务设施 "单位时间内服务完的顾客的平均数", 而今服务机构内有 $n$ 台服务设施, $\dfrac{\lambda}{\mu}$ 是 "单位时间内来到此服务机构的全体顾客所需之总服务时间的平均数", 而今服务机构内共有 $n$ 台服务设施, 所以单位时间内总共能提供 $n$ 个单位时间的服务. 因此, 当 $\dfrac{\lambda}{n\mu} > 1$ 时, 供不应求, 势必等待服务的顾客数愈来愈

多, 对任何 $i \geqslant 0, j \geqslant 0$, 总有

$$\lim_{t\to\infty} P_{i,j}(t) \equiv \lim_{t\to\infty} p(\eta_t = j|\eta_0 = i) = 0,$$

服务系统不能达到平衡状态. 而当 $\dfrac{\lambda}{n\mu} < 1$ 时, 供过于求, 等待服务的顾客数不会愈来愈多, 所以

$$\lim_{t\to\infty} p_{i,j}(t) = \frac{\rho_j}{\sum\limits_{k=0}^{\infty}\rho_k} \quad (i \geqslant 0, j \geqslant 0),$$

亦即服务系统能达到平衡. 令 $p_j = \lim\limits_{t\to\infty} P(\eta_t = j)$, 则

$$p_j = \lim_{t\to\infty}\sum_{i\geqslant 0} P(\eta_0 = i)p_{i,j}(t) = \frac{\rho_j}{\sum\limits_{k=0}^{\infty}\rho_k} \quad (j \geqslant 0),$$

$p_j$ 称服务系统处于平衡状态 $j$ 的概率. 显然

$$\sum_{j=0}^{\infty} p_j = 1,$$

其中 $\rho_k$ 由 (5.35) 所定义, 即是

$$\rho_k = \begin{cases} \left(\dfrac{\lambda}{\mu}\right)^k \dfrac{1}{k!}, & \text{当 } 0 \leqslant k \leqslant n, \\ \left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\cdot\left(\dfrac{\lambda}{n\mu}\right)^{k-n} = \rho_n\left(\dfrac{\lambda}{n\mu}\right)^{k-n}, & \text{当 } k \geqslant n, \end{cases} \tag{5.36}$$

$$\begin{aligned} p_j &= \frac{\rho_j}{\sum\limits_{k=0}^{\infty}\rho_k} = \frac{\rho_j}{\sum\limits_{k=0}^{n-1}\left(\dfrac{\lambda}{\mu}\right)^k \dfrac{1}{k!} + \left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\sum\limits_{k=n}^{\infty}\left(\dfrac{\lambda}{n\mu}\right)^{k-n}} \\ &= \frac{\rho_j}{\sum\limits_{k=0}^{n-1}\left(\dfrac{\lambda}{\mu}\right)^k \dfrac{1}{k!} + \left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\cdot\dfrac{1}{1-\dfrac{\lambda}{n\mu}}} \\ &= \begin{cases} \dfrac{\left(\dfrac{\lambda}{\mu}\right)^j \dfrac{1}{j!}}{\sum\limits_{k=0}^{n-1}\left(\dfrac{\lambda}{\mu}\right)^k \dfrac{1}{k!} + \left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\cdot\dfrac{1}{1-\dfrac{\lambda}{n\mu}}}, & \text{当 } 0 \leqslant j \leqslant n, \\ \dfrac{\left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\cdot\left(\dfrac{\lambda}{n\mu}\right)^{j-n}}{\sum\limits_{k=0}^{n-1}\left(\dfrac{\lambda}{\mu}\right)^k \dfrac{1}{k!} + \left(\dfrac{\lambda}{\mu}\right)^n \dfrac{1}{n!}\cdot\dfrac{1}{1-\dfrac{\lambda}{n\mu}}} = p_n\left(\dfrac{\lambda}{n\mu}\right)^{j-n}, & \text{当 } j \geqslant n. \end{cases} \end{aligned} \tag{5.37}$$

(2) 顾客来到服务机构时, 不能即刻接受服务, 需要等待的概率 $p_w$.

因为服务设施共有 $n$ 台, 所以一个顾客来到服务机构, 不能即刻接受服务的充分必要条件是: 服务系统处于的平衡状态 $j \geqslant n$, 所以

$$p_w = \sum_{j=n}^{\infty} p_j. \tag{5.38}$$

(3) 服务机构内的总顾客 (包括正在服务的与排队等待服务的) 数的均值 $N_m$.

显然

$$N_m = \sum_{j=0}^{\infty} j p_j. \tag{5.39}$$

(4) 等待服务的顾客数的均值 $N_m(W)$.

显然

$$N_m(W) = \sum_{j=0}^{\infty} j p_{n+j}. \tag{5.40}$$

**定理 5.5** 对 $M/M/n$ (等待) 系统而言, 设 $\dfrac{\lambda}{n\mu} < 1$, 当系统达到平衡状态以后 (即 $P(\eta_t = j) = \sum\limits_{i=0}^{\infty} P(\eta_0 = i) p_{i,j}(t) \cong \sum\limits_{i=0}^{\infty} P(\eta_0 = i) p_j = p_j$ (当 $t$ 充分大)), 令来到服务机构请求服务所需之等待时间为 $W$, 则

$$P(W > t) = p_n \bullet \frac{1}{1 - \dfrac{\lambda}{n\mu}} \bullet e^{-(n\mu - \lambda)t} \quad (t \geqslant 0). \tag{5.41}$$

**证** 令 $P_j(W > t) = P(W > t|$ 顾客到达时, 服务系统处于平衡状态 $j)$, 则

$$P(W > t) = \sum_{j=0}^{\infty} p_j P_j(W > t). \tag{5.42}$$

显然

$$P_j(W > t) = 0 \quad (t > 0, j < n). \tag{5.43}$$

将 (5.43) 代入 (5.42) 得

$$P(W > t) = \sum_{j=n}^{\infty} p_j P_j(W > t). \tag{5.44}$$

下面我们来计算 $P_j(W > t)$.

当服务系统处于平衡状态 $j (j \geqslant n)$ 时, 其中 $n$ 个顾客正在接受服务, $j - n$ 个顾客正在排队等待. 所以这时新到来的顾客要在服务机构依次服务完 $j - n + 1$

个顾客后, 他才能被接受服务. 因此若令 $M(t)$ 为时间长度 $t$ 内服务机构服务完的顾客个数, 则

$$P_j(W>t)=\sum_{i=0}^{j-n}P(M(t)=i). \tag{5.45}$$

由于各顾客的服务时间相互独立且具有公共的以 $\mu$ 为参数的负指数分布, 所以在该顾客等待时间内, 每台服务设施的输出过程 (即服务完而离开服务机构的顾客) 是一个以 $\mu$ 为强度的 Poisson 过程 (定理 5.1 正是这一结论的理论根据), 而今该服务机构内共有 $n$ 台设施 (当然它们之间的服务是相互独立的), 所以此服务机构的总的输出过程是一个以 $n\mu$ 为强度的 Poisson 过程. 所以

$$P(M(t)=i)=e^{-n\mu t}\frac{(n\mu t)^i}{i!}. \tag{5.46}$$

将 (5.45) (5.46) 和 (5.37) 代入 (5.44) 得

$$\begin{aligned}
P(W>t)&=\sum_{j=n}^{\infty}p_j\sum_{i=0}^{j-n}e^{-n\mu t}\frac{(n\mu t)^i}{i!}\\
&=\sum_{j=n}^{\infty}p_n\left(\frac{\lambda}{n\mu}\right)^{j-n}e^{-n\mu t}\sum_{i=0}^{j-n}\frac{(n\mu t)^i}{i!}\\
&=p_ne^{-n\mu t}\sum_{k=0}^{\infty}\left(\frac{\lambda}{n\mu}\right)^{k}\sum_{i=0}^{k}\frac{(n\mu t)^i}{i!}\\
&=p_ne^{-n\mu t}\sum_{i=0}^{\infty}\frac{(n\mu t)^i}{i!}\sum_{k=i}^{\infty}\left(\frac{\lambda}{n\mu}\right)^{k}\\
&=p_ne^{-n\mu t}\sum_{i=0}^{\infty}\frac{(n\mu t)^i}{i!}\left(\frac{\lambda}{n\mu}\right)^{i}\frac{1}{1-\dfrac{\lambda}{n\mu}}\\
&=p_ne^{-(n\mu-\lambda)t}\cdot\frac{1}{1-\dfrac{\lambda}{n\mu}},
\end{aligned} \tag{5.47}$$

其中 $p_n$ 由 (5.37) 所定义. 定理证毕.

**推论 5.4** 顾客需要等待的概率为

$$p_w=P(W>0)=p_n\cdot\frac{1}{1-\dfrac{\lambda}{n\mu}}. \tag{5.48}$$

这与 (5.38) 的结论是一致的. 因为由 (5.37) 有

$$p_j=p_n\left(\frac{\lambda}{n\mu}\right)^{j-n}\quad(j\geqslant n),$$

所以由 (5.38) 亦有

$$p_w = \sum_{j=n}^{\infty} p_i = p_n \sum_{j=n}^{\infty} \left(\frac{\lambda}{n\mu}\right)^{j-n} = p_n \cdot \frac{1}{1-\dfrac{\lambda}{n\mu}}.$$

**推论 5.5** 顾客到达服务机构后平均等待时间为

$$\begin{aligned}E(W) &= \int_0^{\infty} \left[t\frac{d}{dt}(1-P(W>t))\right] dt \\ &= \int_0^{\infty} tp_n \cdot \frac{1}{1-\dfrac{\lambda}{n\mu}} \cdot (n\mu-\lambda)e^{-(n\mu-\lambda)t} dt \\ &= \frac{p_n}{1-\dfrac{\lambda}{n\mu}} \cdot \frac{1}{n\mu-\lambda} = p_n \cdot \frac{n\mu}{(n\mu-\lambda)^2} \\ &= p_0 \left(\frac{\lambda}{\mu}\right)^n \cdot \frac{1}{n!} \frac{n\mu}{(n\mu-\lambda)^2}.\end{aligned} \tag{5.49}$$

**定理 5.6** 就 $M/M/1$ (等待) 系统而言, 设 $\dfrac{\lambda}{\mu} < 1$, 当系统达到平衡状态以后, 在长为 $L$ 的时间区间内 ($L$ 相当大).

(1) 忙期 (即服务设施正在服务) 的总长度及平均个数分别为

$$L_B = \frac{\lambda}{\mu}L, \quad Z(L_B) = L\frac{(\mu-\lambda)\lambda}{\mu}. \tag{5.50}$$

(2) 单个忙期的平均长度为

$$L_B(1) = \frac{1}{\mu-\lambda}. \tag{5.51}$$

(3) 单个忙期内所服务的顾客平均数为

$$N(L_B(1)) = \frac{\mu}{\mu-\lambda}. \tag{5.52}$$

**证** 对 $M/M/1$ (等待) 系统而言, 正如定理 5.4 后面, 对 $\dfrac{\lambda}{n\mu} < 1$ (而今 $n=1$) 的意义的分析那样, $\dfrac{\lambda}{\mu}$ 是 "单位时间内来到此服务机构的全体顾客所需之总服务时间之平均数", 亦即在相当长的时间 $L$ 内, 忙期总长度为

$$L_B = \frac{\lambda}{\mu}L.$$

于是, 各闲期 (从服务机构变成闲置开始到新顾客到来为止这段时间) 的总长度为

$$L - L_B = L \cdot \left(1 - \frac{\lambda}{\mu}\right).$$

而在时间 $L$ 内, 闲期的平均个数为

$$\frac{L\left(1-\dfrac{\lambda}{\mu}\right)}{\dfrac{1}{\lambda}}=\lambda L\bullet\left(1-\frac{\lambda}{\mu}\right).$$

由于各个忙期与各个闲期在时间轴上是相间的, 所以, 在时间 $L$ 内, 忙期的平均个数亦为

$$Z(L_B)=\lambda L\bullet\left(1-\frac{\lambda}{\mu}\right)=L\bullet\frac{(\mu-\lambda)\lambda}{\mu}.$$

从而单个忙期的平均长度为

$$L_B(1)=\frac{L_B}{Z(L_B)}=\frac{1}{\mu-\lambda}.$$

单个忙期内所服务的顾客的平均数为

$$N(L_B(1))=\frac{L_B(1)}{\dfrac{1}{\mu}}=\frac{\mu}{\mu-\lambda}$$

(因为在忙期内, 每个顾客的平均服务时间为 $\dfrac{1}{\mu}$) 定理证毕.

(二) $M/M/n$ (消失) 系统

在这一段中, 恒设

(1) 输入过程是以 $\lambda$ 为强度的 Poisson 过程.

(2) 服务规则是消失制, 即当顾客到达时, 若 $n$ 个服务设施至少有一个空闲, 则顾客马上被接受服务; 若 $n$ 个服务设施均被占用, 则顾客就离开服务机构, 或者说此顾客消失了.

(3) 服务机构中服务设施的数目为 $n$, 每个顾客的服务时间服从以 $\mu$ 为参数的负指数分布, 从而每个顾客的平均服务时间为 $\dfrac{1}{\mu}$.

(4) 各顾客的服务时间是相互独立的, 且与输入过程亦相互独立.

设 $\eta_t$ 是时刻 $t$ 正在被服务的顾客数, 则由于现在考虑的是 $M/M/n$ (消失) 系统, 所以 $\eta_t$ 所可能取的值为 $0,1,2,\cdots,n$, 共 $n+1$ 个值.

**定理 5.7** $\{\eta_t : t\in[0,\infty)\}$ 是一个有限状态的 (共 $n+1$ 个状态) 时齐的 Markov 过程, 若令 $P(t)=(p_{i,j}(t),i,j=0,1,\cdots,n)$ 是其转移矩阵, $Q=P'(0)=$

$(p'_{i,j}(0), i, j = 0, 1, \cdots, n)$ 是其转移强度矩阵, 记 $p'_{i,j}(0) = q_{i,j}$, 则

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & & 0 & 0 \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mu_n & -\mu_n \end{pmatrix},$$

其中 $\lambda_i \equiv \lambda (0 \leqslant i \leqslant n), \mu_j = j\mu (1 \leqslant j \leqslant n)$. 而且

$$p_i \equiv \lim_{t\to\infty} P(\eta_t = j) = p_0 \cdot \frac{\left(\dfrac{\lambda}{\mu}\right)^j}{j!} \quad (j = 1, \cdots, n), \tag{5.53}$$

其中

$$p_0 = \left(\sum_{j=0}^{n} \frac{\left(\dfrac{\lambda}{\mu}\right)^j}{j!}\right)^{-1} \tag{5.54}$$

(5.53), (5.54) 通常称为 **Erlang 公式**. 它给出了 $t \to \infty$ 时, 系统处于平衡状态 $j$ 的概率, 特别地

$$p_n = \left(\frac{\left(\dfrac{\lambda}{\mu}\right)^n}{n!}\right)\left(\sum_{j=0}^{n} \frac{\left(\dfrac{\lambda}{\mu}\right)^j}{j!}\right)^{-1} \tag{5.55}$$

是 $n$ 个服务设施都正在进行服务的概率, 亦即顾客到达时遭到拒绝服务而消失的概率, 称之为**消失概率**.

由 (5.53), (5.54) 得知: 正在服务机构内被服务的顾客的平均数为

$$N_m = \sum_{j=0}^{n} j p_j = p_0 \sum_{j=0}^{n} j \frac{\left(\dfrac{\lambda}{\mu}\right)^j}{j!} = \frac{\lambda}{\mu} p_0 \sum_{j=0}^{n-1} \frac{\left(\dfrac{\lambda}{\mu}\right)^j}{j!}. \tag{5.56}$$

**证** 仿例 5.1 知 $\{\eta_t\}$ 是时齐的 (有限状态的) Markov 过程. 又

$$\begin{aligned} & P(\eta_{t+\Delta t} = i+1 | \eta_t = i) \\ = \ & P([t, t+\Delta t) \text{ 内恰来一个顾客, 且 } i \text{ 个正在服务的顾客} \\ & \text{尚无一个结束服务}) \\ & + P([t, t+\Delta t) \text{ 内至少来两个顾客, 且使 } \eta_{t+\Delta t} = i+1) \\ \stackrel{\text{记作}}{=\!=} \ & \mathrm{I}(\Delta t) + \mathrm{II}(\Delta t). \end{aligned} \tag{5.57}$$

根据 Poisson 输入与负指数分布的性质, 以及输入过程与各顾客的服务时间相互独立可知:

$$
\begin{aligned}
\mathrm{I}(\Delta t) &= (\lambda \Delta t e^{-\lambda \Delta t})(e^{-\mu \Delta t})^i \\
&= \lambda \Delta t + o(\Delta t) \quad (\Delta t \to 0), \qquad (5.58)\\
\mathrm{II}(\Delta t) &= o(\Delta t) \quad (\Delta t \to 0). \qquad (5.59)
\end{aligned}
$$

将 (5.58), (5.59) 代入 (5.57) 得

$$
q_{i,i+1} = \lim_{\Delta t \to 0} \frac{\mathrm{I}(\Delta t) + \mathrm{II}(\Delta t)}{\Delta t} = \lambda \quad (n > i \geqslant 0), \qquad (5.60)
$$

而

$$
\begin{aligned}
& P(\eta_{t+\Delta t} = i-1 | \eta_t = i) \\
= \ & P(\text{在 } [t, t+\Delta t) \text{内无顾客到来, 且 } i \text{ 个顾客中恰有一个服务结束}) \\
& + P(\text{在 } [t, t+\Delta t) \text{ 内至少来一个顾客, 且 } i \text{ 个顾客中} \\
& \text{至少有一个顾客服务结束, } \eta_{t+\Delta t} = i-1) \\
\stackrel{\text{记作}}{=\!=} \ & \mathrm{I}^*(\Delta t) + \mathrm{II}^*(\Delta t). \qquad (5.61)
\end{aligned}
$$

而

$$
\begin{aligned}
\mathrm{I}^*(\Delta t) &= e^{-\lambda \Delta t} i (1 - e^{-\mu \Delta t})(e^{-\mu \Delta t})^{i-1} \\
&= i\mu \Delta t + o(\Delta t) \quad (\Delta t \to 0), \qquad (5.62)\\
0 \leqslant \mathrm{II}^*(\Delta t) &\leqslant (1 - e^{-\lambda \Delta t})(1 - (e^{-\mu \Delta t})^i) \\
&= o(\Delta t) \quad (\Delta t \to 0), \qquad (5.63)
\end{aligned}
$$

所以

$$
q_{i,i-1} = \lim_{\Delta t \to 0} \frac{\mathrm{I}^*(\Delta t) + \mathrm{II}^*(\Delta t)}{\Delta t} = i\mu \quad (i \geqslant 1). \qquad (5.64)
$$

仿之易证:

$$
\begin{cases}
q_{0,0} = -\lambda, \\
q_{i,i} = -(\lambda + i\mu)(n > i \geqslant 1), \quad q_{n,n} = -n\mu, \\
q_{i,j} = 0 \quad (|i - j| > 1).
\end{cases} \qquad (5.65)
$$

总之 $Q$ 之值如定理 5.7 所断言. 下证 (5.53), (5.54).

再用本章 §6 习题 2 (6), 解线性方程组:

$$\begin{cases} (x_0, x_1, \cdots, x_n) \begin{pmatrix} q_{0,0} & q_{0,1} & & q_{0,n} \\ q_{1,0} & q_{1,1} & & q_{1,n} \\ \vdots & \vdots & & \vdots \\ q_{n,0} & q_{n,1} & \cdots & q_{n,n} \end{pmatrix} = 0, \\ x_i \geqslant 0 \quad (0 \leqslant i \leqslant n), \end{cases}$$

得通解为

$$x_j = \frac{\lambda_0 \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \mu_2 \cdots \mu_i} x_0 = \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j x_0 \quad (j = 1, 2, \cdots, n).$$

所以, 由本章 §6 习题 2 (6), 得

$$\lim_{t \to \infty} p_{i,j}(t) = \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j \left(\sum_{j=0}^{n} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j\right)^{-1}$$

$(i = 0, 1, \cdots, n, j = 0, 1, \cdots, n)$, 故

$$\begin{aligned} p_j &\equiv \lim_{t \to \infty} P(\eta_t = j) \\ &= \lim_{t \to \infty} \sum_{i=0}^{n} P(\eta_0 = i) p_{i,j}(t) \\ &= \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j \left(\sum_{j=0}^{n} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j\right)^{-1} \quad (0 \leqslant j \leqslant n). \end{aligned}$$

定理证毕.

(三) $M/M/\infty$ 系统

在这一段中, 恒设

(1) 输入过程是以 $\lambda$ 为强度的 Poisson 过程.

(2) 服务规则就是先到先服务, 注意, 由于现在服务设施有无穷多个 (理想的), 故任一来到的顾客, 既不需等待也不消失, 因此这是无等待制与消失制的差别.

(3) 服务机构中服务设施有无穷多个, 每个顾客的服务时间服从以 $\mu$ 为参数的负指数分布, 从而每个顾客的平均服务时间为 $\dfrac{1}{\mu}$.

(4) 各顾客的服务时间是相互独立的, 且与输入过程亦相互独立.

设 $\eta_t$ 是时刻 $t$ 正在服务的顾客数.

例 5.1 已证 $\{\eta_t : t \in [0,\infty)\}$ 是一个生灭过程, 其转移强度矩阵由 (5.19) 所给出, 且其中 $\lambda_i \equiv \lambda, \mu_i = i\mu$. 由于生灭过程 $\{\eta_t : t \in [0,\infty)\}$ 的转移强度矩阵 $Q = (q_{i,j}, i, j \geqslant 0)$ 已求出, $\{\eta_t : t \in [0,\infty)\}$ 的其他概率性质化为由 $Q$ 来研究其他相应的性质. 这是纯数学的问题了. 在此我们不准备再深入讨论.

## §6 习题及应用

在 §2 中, 我们系统地研究了时齐的可数状态的 Markov 转移矩阵 $P(t) = (p_{i,j}(t), i, j \in E)$ 的连续性、可微性以及遍历性. 这一节的习题主要涉及两方面: 一是当状态空间 $E$ 是有限时, §2、§3 中相应的结果是否更简化、更强? 二是对准转矩阵来说, 转移矩阵的一些结果是否还保持.

先考虑第一方面的问题. 习题 1 和 2 中的状态空间 $E$ 都是有限集.

1. 设 $P(t) = (p_{i,j}(t), i, j \in E)$ 是一个标准的转移矩阵, 试证恒有:

(1) $P'(0) = \lim\limits_{t\to 0+} \dfrac{1}{t}(P(t) - I)$

存在, 记此极限为 $Q, Q$ 是保守的转移强度矩阵;

(2) $P'(t) = QP(t) = P(t)Q (\forall t \geqslant 0)$;

(3) $P(t) = e^{Qt}$.

反之. 任给一个保守的转移强度矩阵 $Q$, 则 $P(t) \stackrel{\text{def.}}{=\!=} e^{Qt}$ 是一个标准的转移矩阵.

2. 设 $P(t) = (p_{i,j}(t), i, j \in E)$ 是一个标准的转移矩阵, $Q = P'(0)$. 试证

(1) $\lim\limits_{t\to\infty} P(t)$ 存在, 记此极限为 $\Pi$;

(2) $P(t)\Pi = \Pi P(t) = \Pi^2 = \Pi (\forall t \geqslant 0)$;

(3) $\Pi \geqslant 0, \Pi\mathbf{1} = \mathbf{1}$;

(4) $\Pi Q = Q\Pi = 0$;

(5) $\Pi = \mathbf{1}\pi', \pi' > 0$ (此处 $\pi' = (\pi_i, i \in E)$ 是 $E$ 上的一个概率分布, 每项均大于 0) 的充分必要条件是: 存在 $t_0 > 0$, 使 $P(t_0) > 0$;

(6) $\Pi = \mathbf{1}\pi' \Leftrightarrow$《$x'Q = 0, x' \geqslant 0,$》的通解为 $c\bar{x}'$, 其中 $c$ 为任一非负常数, $\bar{x}' \neq 0$, 这时 $\pi' = c\bar{x}'/(c\bar{x}')\mathbf{1}$.

现在再考虑标准的准转移矩阵的分析理论. 习题 3~10 中的 $P(t) = (p_{i,j}(t), i, j \in E)$ 是标准的准转移矩阵, $E$ 是任意可数集.

3. 试证对上述 $P(t) = (p_{i,j}(t), i, j \in E)$ 恒有

$$|p_{i,J}(t) - p_{i,J}(s)| \leqslant 1 - p_{i,i}(|s-t|),$$

其中 $i \in E, J \subset E, s, t \geqslant 0, p_{i,J}(t) = \sum\limits_{j\in J} p_{i,j}(t)$.

4. 试证上述 $P(t)$ 恒有 $\lim\limits_{t\to\infty} P(t) = \Pi$ 存在, 记此极限为 $\Pi$, 则还有

$$\Pi P(t) = P(t)\Pi = \Pi^2 = \Pi.$$

5. 试证对上述 $P(t)=(p_{i,j}(t),i,j\in E)$ 总有

$$\lim_{t\to 0+}\frac{1}{t}(1-p_{i,i}(t)) \text{ 存在 (可以为 } \infty),$$

记此极限为 $q_i$ (或 $-q_{i,i}$), 则还有

$$0\leqslant q_i\leqslant\infty,\quad p_{i,i}(t)\geqslant e^{-q_it}\quad (i\in E,t\geqslant 0,e^{-\infty}\text{ 定义为 }0).$$

特别地, 当 $q_i=0$ 时, $p_{i,i}(t)\equiv 1, p_{i,j}(t)\equiv 0(\forall j\neq i)$.

6. 令

$$\mathscr{G}=\left\{J:J\subset E,\lim_{t\to 0+}\sup_{j\in J}(1-p_{j,j}(t))=0\right\}.$$

试证: 对任意固定的 $i\in E,J\in\mathscr{G},i\bar{\in}J$, 有

(1) 当 $K\subset J$ 时 $\lim\limits_{t\to 0+}\dfrac{p_{i,K}(t)}{t}$ 存在, 记此极限为 $q_{i,K}$, 则 $0\leqslant q_{i,K}<\infty$ (其中 $p_{i,K}(t)=\sum\limits_{k\in K}p_{i,K}(t)$), 而且此极限对 $K\subset J$ 还一致成立. 此外, $q_{i,K}$ 对 $K\subset J$ 来说, 还有完全可加性.

(2) 特别地, 当 $J=\{j\}$ 是单点集, $i,j\in E,i\neq j$, 有

$$\lim_{t\to 0+}\frac{p_{i,j}(t)}{t}=q_{i,j}\text{ 存在},$$

且 $0\leqslant q_{i,j}<\infty,\sum\limits_{j\neq i}q_{i,j}\leqslant q_i$.

7. 设 $P(t)=(p_{i,j}(t),i,j\in E),q_i$ 如上定义, 试证下列二条件等价:

(1) $\sup\limits_{i\in E}q_i<\infty$;

(2) $\lim\limits_{t\to\infty}\sup\limits_{i\in E}(1-p_{i,i}(t))=0$.

如果上述条件中有一个成立, 则

$$\lim_{t\to 0+}\frac{1-p_{i,i}(t)}{t}=q_i\quad\text{对 } i\in E \text{ 一致成立},$$

而且还有:

$$\sum_{j\in E-\{i\}}q_{i,j}=\lim_{t\to 0+}\sum_{j\in E-\{i\}}\frac{p_{i,j}(t)}{t}\quad (i\in E).$$

8. 设 $P(t)=(p_{i,j}(t),i,j\in E),q_i,q_{i,j}$ 如前定义. 固定 $i\in E$, 设 $0<q_i<\infty$, 则有唯一一组 $g_{i,j}(t)(j\in E,t\geqslant 0),g_{i,j}(\cdot)$ 在 $[0,\infty)$ 上连续, 使

$$p_{i,j}(t)=q_ie^{-q_it}\int_0^t e^{q_it}g_{i,j}(s)ds+e^{-q_it}\delta_{i,j}\quad (j\in E,t\geqslant 0).$$

此外, 还有

(1) $g_{i,j}(t) \geqslant 0(\forall j \in E, t \geqslant 0), \sum\limits_{j\in E} g_{i,j}(t) \leqslant 1(t > 0)$;

(2) $g_{i,j}(s+t) = \sum\limits_{k\in E} g_{i,k}(s)p_{k,j}(t)(s > 0, t \geqslant 0, j \in E)$;

(3) $\lim\limits_{t\to 0+} g_{i,j}(t) = \begin{cases} 0, & 当\ j = i, j \in E, \\ q_{i,j}/q_i, & 当\ j \neq i, j \in E; \end{cases}$

(4) $\lim\limits_{t\to\infty} g_{i,j}(t) = \lim\limits_{t\to\infty} p_{i,j}(t) = \pi_{i,j}(j \in E)$,

因此, $p_{i,j}(t)$ 在 $(0,\infty)$ 内有连续微商 $p'_{i,j}(t)$, 它可以表为:

$$p'_{i,j}(t) = q_i g_{i,j}(t) - q_i p_{i,j}(t).$$

由此式及 $g_{i,j}(t)$ 的性质还可以得出 $p'_{i,j}(t)$ 满足:

(1) $\sum\limits_{j\in E} |p'_{i,j}(t)| \leqslant 2q_i \quad (t \geqslant 0)$;

(2) $\sum\limits_{j\in E} p'_{i,j}(t) \leqslant 0 \quad (t \geqslant 0)$;

(3) $p'_{i,j}(s+t) = \sum\limits_{k\in E} p'_{i,k}(s)p_{k,j}(t) \quad (s > 0, t \geqslant 0, j \in E)$;

(4) $\lim\limits_{t\to 0+} p'_{i,j}(t) = p'_{i,j}(0) = q_{i,j} \quad (j \in E)$;

(5) $\lim\limits_{t\to\infty} p'_{i,j}(t) = 0 \quad (j \in E)$.

9. 设 $P(t) = (p_{i,j}(t), i, j \in E), q_i, q_{i,j}$ 如前定义. 固定 $i \in E$, 设 $q_i < \infty$, 则下列二陈述等价:

(1) $p'_{i,j}(t) = \sum\limits_{k\in E} q_{i,k}p_{k,j}(t) \quad (t \geqslant 0, j \in E)$,

(2) $\sum\limits_{j\in E-\{i\}} q_{i,j} = \lim\limits_{t\to 0+} \sum\limits_{j\in E-\{i\}} \dfrac{p_{i,j}(t)}{t}$.

10. 设 $P(t), q_i$ 如前所定义, 假定 $\sup\limits_{i\in E} q_i < \infty, P'(0) = Q$. 试证

(1) $P'(t) = QP(t) = P(t)Q \quad (t \geqslant 0)$;

(2) $P(t) = e^{Qt} \quad (t \geqslant 0)$;

(3) $\varPi Q = Q\varPi = 0$, 其中 $\varPi = \lim\limits_{t\to\infty} P(t)$.

反之, 任意给定一个转移强度矩阵 $Q = (q_{i,j}, i, j \in E)$, 若 $\sup\limits_{i\in E}(-q_{i,i}) < \infty$, 则 $P(t) \stackrel{\text{def.}}{=\!=} e^{Qt}$ 是一个标准的准转移矩阵.

11. 证明定理 2.1.

# *第八章　随机环境中的 Markov 链

我们知道, 随机徘徊是 Markov 链中最简单的非常典型的一个特例. 它在物理学、渗透理论、统计中的序贯分析, 都有很重要的应用. 在讲随机环境中的 Markov 链时, 我们也用随机环境中的随机徘徊作为引导, 来说明随机环境中的 Markov 链的基本思想.

我们从最简单的随机徘徊 —— 经典的 Bernoulli 随机徘徊开始.

所谓经典的 Bernoulli 随机徘徊 , 就是定义在概率空间 $(\Omega,\mathscr{F},P)$ 上的取整数值的 Markov 链 $\{S_n,n\geqslant 0\}$, 满足

$$P(S_{n+1}=x+1|S_n=x)=q,$$
$$P(S_{n+1}=x-1|S_n=x)=p\overset{\text{def.}}{=\!=}1-q,$$

(其中 $n\geqslant 0,x\in\mathbf{Z}=\{0,\pm1,\pm2,\cdots\},q\in(0,1)$.) 对于这种随机徘徊, 其概率规律性完全决定于其转移概率 $q$.

如果存在一些随机因素, 影响其转移概率 $q$, 而且这种随机干扰还依赖于时间 $n$ 和正在徘徊的 "质点" 所处的位置 $x$, 那么, $q$ 将被一个随机场 $\{q(n,x,\omega),n\geqslant 0,x\in\mathbf{Z}\}$ 所代替. 为简单起见, 我们记 $q(n,x,\bullet)=q(n,x)$, 记此随机场为 $q=\{q(n,x),n\geqslant 0,x\in\mathbf{Z}\}$. 这时经典的 Bernoulli 随机徘徊就变成了 "**依时依空的随机环境中的 Bernoulli 随机徘徊**". 它的正式定义是:

设 $(\Omega,\mathscr{F},P)$ 是一个概率空间, $\{S_n,n\geqslant 0\}$ 是其上的取值于 $\mathbf{Z}$ 的一列随机变量, $q=\{q_{n,x},n\geqslant 0,x\in\mathbf{Z}\}$ 是 $(\Omega,\mathscr{F},P)$ 上取值于 (0,1) 的随机场, 记 $P_q(\bullet)=P(\bullet|q)$ 为条件概率. 不妨设 $S_0=0$. 若 $\{S_n,n\geqslant 0\}$ 对概率测度 $P_q$ 是

一个 Markov 链 (未必时齐), 且有下述转移概率:

$$P_q\left(S_{n+1}=x+1|S_n=x\right)=q(n,x),$$
$$P_q\left(S_{n+1}=x-1|S_n=x\right)=1-q(n,x),$$

则称 $\{S_n,n\geqslant 0\}$ 是一个依时依空的随机环境中的 Bernoulli 随机徘徊, $q=\{q(n,x),n\geqslant 0,x\in\mathbf{Z}\}$ 称为其环境场. 特别地, 若 $q(n,x)$ 不依赖 $x$, 则称 $\{S_n,n\geqslant 0\}$ 是**依时的随机环境中的 Bernoulli 随机徘徊**; 类似地, 若 $q(n,x)$ 不依赖 $n$, 则称 $\{S_n,n\geqslant 0\}$ 为 **依空的随机环境中的 Bernoulli 随机徘徊**.

Berard J. 在 2004 年的 "J.Appl.Prob.41 卷 P.83 – 92" 上证明了下面的 a.s. 中心极限定理:

若 $\{S_n,n\geqslant 0\}$ 是上面定义的依时依空的随机环境中的 Bernoulli 随机徘徊, 而且其环境场 $q=\{q(n,x),n\geqslant 0,x\in\mathbf{Z}\}$ 是独立同分布的且 $\boldsymbol{E}(q(n,x))=\dfrac{1}{2}$ (对一切 $n\geqslant 0,\ x\in\mathbf{Z}$), 则

$$\lim_{n\to\infty}P_q\left(\frac{S_n}{\sqrt{n}}<\alpha\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\alpha}e^{-\frac{t^2}{2}}dt,\ \text{a.s.},$$

注意上式左边是关于 $\mathscr{F}$ 的子 $\sigma$ 代数 $\sigma(q)\overset{\text{def.}}{=\!=}\sigma\{q(n,x),n\geqslant 0,x\in\mathbf{Z}\}$ 可测的随机变量, 故上式是 a.s. 成立.

如果环境场 $q$ 是一个 (0,1) 中的常数, 即 $q(n,x,\omega)=$ 常数 $q$ (不依赖 $n,x,\omega$), 则上述中心极限定理就是古典的最简单的中心极限定理 —— Bernoulli 中心极限定理.

从上面的简单例子看出, 随机环境中的 Markov 链, 要比经典的 Markov 链复杂得多. 它涉及两个随机过程 (或一个随机过程和另一个随机场) 的耦合. 它的转移概率不再是一个 [0,1] 中的数, 而是一个取值于 [0,1] 的随机函数. 从上面的定理也可以看出, 一个经典随机徘徊中的最简单的中心极限定理, 2004 年还有人研究.

本章主要讨论的是依时的随机环境中的 Markov 链. 依空的随机环境中的 Markov 链请参看文献 [119], 那里主要研究依空的随机环境中的随机徘徊. 依时依空的随机环境中的 Markov 链的研究还不多, 本章只作一点简单介绍. 至于连续时间参数的随机环境中的 Markov 过程, 本章完全不涉及了. 有兴趣的读者可参看文献 [60]、[61]、[79].

## §1 依时随机环境中的 Markov 链的基本概念及存在性

在这一节中, 恒令 $\mathbf{Z}=\{0,\pm1,\pm2,\cdots\}$ 为整数集, $\mathbf{N}\overset{\text{def.}}{=\!=}\mathbf{Z}_+$ 为非负整数集, $\mathbf{R}$ 是实数集. 令 $\mathscr{X}$ 是一个可数集, $\mathscr{A}$ 是 $\mathscr{X}$ 中的离散 $\sigma$ 代数, 即是 $\mathscr{X}$ 中的一

切子集所成的 $\sigma$ 代数, 再令 $(\Theta,\mathscr{B})$ 是任意的一个可测空间.

按常规符号,

$\vec{\mathscr{X}} \stackrel{\text{def.}}{=} \mathscr{X}^{\mathbf{N}}$ 是乘积空间, $\vec{\mathscr{A}} \stackrel{\text{def.}}{=} \mathscr{A}^{\mathbf{N}}$ 是乘积 $\sigma$ 代数. 类似地, 再记 $\vec{\Theta} \stackrel{\text{def.}}{=} \Theta^{\mathbf{Z}}, \vec{\mathscr{B}} \stackrel{\text{def.}}{=} \mathscr{B}^{\mathbf{Z}}$. $\mathscr{X},\Theta,\vec{\mathscr{X}},\vec{\Theta}$ 中的元素分别用 $x,\theta,\vec{x}$ 和 $\vec{\theta}$ 记之. 记 $T$ 是 $\vec{\Theta}$ 到 $\vec{\Theta}$ 的推移算子, 即

$$(T\vec{\theta})_n = \theta_{n+1} \quad (\forall \vec{\theta} = (\theta_n, n\in \mathbf{Z})).$$

$\mathbf{1}_A(\bullet) = \delta(\bullet, A)$ 表示集合 $A$ 上的示性函数.

**定义 1.1**　设 $M(\mathscr{X},\mathscr{A})$ 表 $\mathscr{X}\times\mathscr{X}$ 上的一切转移矩阵 (转移矩阵之定义见第六章定义 2.1),

$$p(\bullet): \Theta \mapsto M(\mathscr{X},\mathscr{A}),$$

若对每一对 $x,y\in\mathscr{X}$, $p(\theta)(x,y)$ 都是 $\theta$ 的关于 $\sigma$ 代数 $\mathscr{B}$ 可测函数, 则称 $p(\bullet)$ 是一个**随机转移矩阵**. 记 $p(\theta)(x,y)$ 为 $p(\theta;x,y)$.

注意: 在第六章的确定 (非随机) 环境中的 Markov 链中, 其转移矩阵写作 $P=(p_{x,y},x,y\in\mathscr{X})$, 在本章随机环境中的 Markov 链的随机转移矩阵, 我们用 $p$ 或 $p(\bullet)$ 表示, $p(\theta;x,y)\stackrel{\text{def.}}{=}p(\theta)(x,y)$ 是此随机转移矩阵在固定随机参数 $\theta$ 下对应于 $(x,y)$ 的元素.

我们记

$$\begin{aligned}
&p^{(n)}(\vec{\theta}) \stackrel{\text{def.}}{=} p(\theta_0,\theta_1,\cdots,\theta_{n-1})\\
&\stackrel{\text{def.}}{=} p(\theta_0)p(\theta_1)\cdots p(\theta_{n-1}) \quad (\forall\vec{\theta}=(\theta_n,n\in\mathbf{Z})\in\vec{\Theta})
\end{aligned} \tag{1.1}$$

为 $n$ 个转移矩阵的乘积 (注意: 固定 $\vec{\theta}\in\vec{\Theta}, p(\theta_i)$ 是一个转移矩阵), 并记

$$\begin{aligned}
&p^{(n)}(\vec{\theta})(x,y) \stackrel{\text{def.}}{=} p^{(n)}(\vec{\theta};x,y)\\
&\stackrel{\text{def.}}{=} p(\theta_0,\theta_1,\cdots,\theta_{n-1})(x,y) \stackrel{\text{def.}}{=} p(\theta_0,\theta_1,\cdots,\theta_{n-1};x,y)\\
&\stackrel{\text{def.}}{=} (p(\theta_0)p(\theta_1)p(\theta_{n-1}))(x,y),\ (\forall x,y\in\mathscr{X}),
\end{aligned} \tag{1.2}$$

即上述五个符号都表示 (1.1) 式右方 $n$ 个矩阵之积在 $(x,y)$ 之元素.

记全体随机转移矩阵为 $RM(\Theta;\mathscr{X},\mathscr{A})$.

**定义 1.2**　设 $(E,\mathscr{E})$ 是任一可测空间, 令函数族 $\{P^{(n)}(\bullet,\bullet),n\geqslant 1\}$ 满足:

(1) 固定任何 $n\geqslant 1, P^{(n)}(\bullet,\bullet): E\times\mathscr{E}\mapsto[0,1]$;

(2) 固定 $n\geqslant 1, x\in E, P^{(n)}(x,\bullet)$ 是 $\mathscr{E}$ 上的概率测度;

(3) 固定 $n\geqslant 1, A\in\mathscr{E}$, $P^{(n)}(\bullet,A)$ 关于 $\mathscr{E}$ 可测;

(4) 满足 $(K-c)$ 方程式: 对任何 $m,n\geqslant 1, x\in E, A\in\mathscr{E}$ 有

$$P^{(m+n)}(x,A) = \int_E P^{(m)}(x,dy)P^{(n)}(y,A). \tag{1.3}$$

则称 $P^{(n)}(x,A)$ 为 ($(E,\mathscr{E})$ 上的) **转移函数**. 记 $P(x,A)=P^{(1)}(x,A)$.

**定义 1.3** 设 $\{X_n, n=0,1,\cdots\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以任意可测空间 $(E,\mathscr{E})$ 为状态空间的随机过程, $\mu$ 是 $\mathscr{E}$ 上的概率测度, $P^{(n)}(x,A)(n\geqslant 1, x\in E, A\in\mathscr{E})$ 是转移函数, 如果对任何 $\{A_0,\cdots,A_n\}\subset\mathscr{E}$, 都有

$$\begin{aligned}&P\left(\bigcap_{i=0}^{n}\{X_i\in A_i\}\right)\\&=\int_{A_0}\mu(dx_0)\int_{A_1}P(x_0,dx_1)\cdots\int_{A_n}P(x_{n-1},dx_n),\end{aligned}\tag{1.4}$$

则称 $\{X_n, n=0,1,\cdots\}$ 为以 $\mu$ 为初始分布, 以 $(E,\mathscr{E})$ 为状态空间, 以 $P^{(n)}(x,A)$ 为转移函数的**时齐的 Markov 链**.

**命题 1.1** 设 $\{X_n, n\geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的以 $(E,\mathscr{E})$ 为状态空间的以 $P^{(n)}(x,A)$ 为转移函数的时齐的 Markov 链, 则对任何 $n\geqslant 1, m\geqslant 0, A_i\in\mathscr{E}(0\leqslant i\leqslant m+n, i\neq m), x\in E$, 总有

$$\begin{aligned}&P(X_{m+n}\in A_{m+n}|X_m=x, X_k\in A_k, 0\leqslant k<m)\\&=P(X_{m+n}\in A_{m+n}|X_m=x)\\&=P^{(n)}(x,A_{m+n}).\end{aligned}\tag{1.4$^*$}$$

(当左方有意义时.)

证明甚易, 读者可作为习题验证之.

**附注 1.1** 方程式 (1.4) 和 (1.4)$^*$ 都称为 "Markov 性".

**定义 1.4** 设 $\overrightarrow{X}=\{X_n, n\in\mathbf{N}\}$ 和 $\overrightarrow{\xi}=\{\xi_n, n\in\mathbf{Z}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的两个随机变量列. $X_n:\Omega\mapsto\mathscr{X}, \xi_n:\Omega\mapsto\Theta, p(\cdot)\in RM(\Theta;\mathscr{X},\mathscr{A})$. 如果下面的条件成立:

$$\begin{aligned}\text{(M)}\qquad &P\left(X_{n+1}=y, \xi_{n+1}\in B\,\middle|\,\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n\right)\\&=P(\xi_n;X_n,y)P(\xi_{n+1}\in B\,|\,\overrightarrow{\xi}_{-\infty}^n), \text{a.s.}\end{aligned}\tag{1.5}$$

对每个 $y\in\mathscr{X}, B\in\mathscr{B}, n\geqslant 0$. 此处 (今后亦如是) $\overrightarrow{\xi}_k^m\overset{\text{def.}}{=\!=}\{\xi_n, k-1<n<m+1\}, -\infty\leqslant k\leqslant m\leqslant\infty, \overrightarrow{X}_0^n\overset{\text{def.}}{=\!=}\{X_0,X_1,\cdots,X_n\}, n\geqslant 0$. 则称 $(\overrightarrow{X},\overrightarrow{\xi})$ 是一个具有随机转移矩阵 $p$ 的**依时随机环境中的 Markov 链** (简记为 MCTRE). 称 $\overrightarrow{X}$ 为本原链, $\overrightarrow{\xi}$ 为**环境链** (或环境).

注意: (1.5) 式中 $\overrightarrow{X}_0^n$ 和 $\overrightarrow{\xi}_{-\infty}^n$ 都是随机变量列, (1.5) 两边是关于它们所产生的 $\sigma$ 代数的条件概率.

**定义 1.5** 概率空间 $(\Omega^*, \mathscr{F}^*, P^*)$ 上取值于 $\mathscr{X}$ 的随机变量序列 $\{Y_n, n \geqslant 0\}$ 称为一个 $p - \overrightarrow{\theta}$ 链, 如果

$$P^*\left(\bigcap_{i=0}^{n}\{Y_i = x_i\}\right) = \varphi(x_0, \overrightarrow{\theta})\prod_{i=0}^{n} p(\theta_i; x_i, x_{i+1}), \tag{1.6}$$

对任何 $x_0, x_1, \cdots, x_n \in \mathscr{X}, n \geqslant 0$ 成立, 其中 $\overrightarrow{\theta} \in \overrightarrow{\Theta}, p(\bullet) \in RM(\Theta; x, \mathscr{A})$, $\varphi(\bullet, \overrightarrow{\theta})$ 是 $\mathscr{X}$ 上的概率分布.

显然 $p - \overrightarrow{\theta}$ 链是非时齐的 Markov 链.

**定义 1.6** 设 $p(\bullet) \in RM(\Theta; \mathscr{X}, \mathscr{A}), \Phi$ 是 $\mathscr{A} \times \overrightarrow{\mathscr{B}}$ 上的一个概率测度. 我们称 $V = (\mathscr{X}, \mathscr{A}; \Theta, \mathscr{B}; p, \Phi)$ 是一个 $p - \Phi$ 链.

令

$$L\begin{pmatrix} 0, & 1, & \cdots, & n \\ x_0, & x_1, & \cdots, & x_n \end{pmatrix} = \left\{\overrightarrow{x} \in \overrightarrow{\mathscr{X}} : (\overrightarrow{x})_i = x_i, 0 \leqslant i \leqslant n\right\},$$

$$L\begin{pmatrix} k, & k+1, & \cdots, & l \\ B_k, & B_{k+1}, & \cdots, & B_l \end{pmatrix} = \left\{\overrightarrow{\theta} \in \overrightarrow{\Theta} : (\overrightarrow{\theta})_i \in B_i, k \leqslant i \leqslant l\right\} \quad (B_i \in \mathscr{B})$$

为可测柱集, 简记

$$L\begin{pmatrix} 0, & 1, & \cdots, & n \\ x_0, & x_1, & \cdots, & x_n \end{pmatrix} = L(\overrightarrow{\mathscr{X}} : x_0, x_1, \cdots, x_n),$$

$$L\begin{pmatrix} k, & k+1, & \cdots, & l \\ B_k, & B_{k+1}, & \cdots, & B_l \end{pmatrix} = L(\overrightarrow{\Theta} : B_k, B_{k+1}, \cdots, B_l).$$

对任意的 $p(\bullet) \in RM(\Theta; \mathscr{X}, \mathscr{A})$, 定义

$$\begin{aligned} & P^x_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}}; x_0, x_1, \cdots, x_n)) \\ = {} & \delta_{x,x_0}\prod_{i=0}^{n-1} p(\theta_i; x_i, x_{i+1}) \ (x \in \mathscr{X}, \overrightarrow{\theta} \in \overrightarrow{\Theta}), \end{aligned} \tag{1.7}$$

容易证明 $P^x_{\overrightarrow{\theta}}(\bullet)$ 可以扩展到 $\overrightarrow{\mathscr{A}}$ 上去而得一概率测度, 仍用 $P^x_{\overrightarrow{\theta}}(\bullet)$ 记此概率测度.

**命题 1.2** 对任何 $A \in \overrightarrow{\mathscr{A}}, x \in \mathscr{X}$, (1.7) 所定义的 $P^x_{\overrightarrow{\theta}}(A)$ 是 $\overrightarrow{\theta}$ 的 $\overrightarrow{\mathscr{B}}$ 可测函数.

证明甚易, 读者可作为习题验证之.

**定义 1.7** 设 $p \in RM(\Theta; \mathscr{X}, \mathscr{A})$ $T$ 是由 $\overrightarrow{\Theta}$ 到 $\overrightarrow{\Theta}$ 的推移算子, $p^{(n)}(\overrightarrow{\theta}; x, y)$ 如 (1.2) 所定义. 对每个正整数 $n \geqslant 1$, 在 $(\mathscr{X} \times \overrightarrow{\Theta}, \mathscr{A} \times \overrightarrow{\mathscr{B}})$ 上定义函数 $\widehat{P}^{(n)}((x, \overrightarrow{\theta}), F)$ 如下:

$$\widehat{P}^{(n)}((x, \overrightarrow{\theta}), F) = \sum_{y \in \mathscr{X}} p^{(n)}(\overrightarrow{\theta}; x, y) \mathbf{1}_{(F)_y}(T^n \overrightarrow{\theta}), \tag{1.8}$$

其中 $(x, \overrightarrow{\theta}) \in \mathscr{X} \times \overrightarrow{\Theta}, F \in \mathscr{A} \times \overrightarrow{\mathscr{B}}, n \geqslant 1, (F)_y = \{\overrightarrow{\theta} \in \overrightarrow{\Theta} : (y, \overrightarrow{\theta}) \in F\}$ 是 $F$ 在 $y$ 处的截口集. 易证:

$$\widehat{P}^{(n)}((x, \overrightarrow{\theta}), \{y\} \times B) = p^{(n)}(\overrightarrow{\theta}; x, y) \mathbf{1}_B(T^n \overrightarrow{\theta}). \tag{1.8$^*$}$$

记 $\widehat{P}((x, \theta), F) = \widehat{P}^{(1)}((x, \overrightarrow{\theta}), F)$.

**命题 1.3** (1.8) 式中所定义的 $\widehat{P}^{(n)}((x, \overrightarrow{\theta}), F)$ 是 $(\mathscr{X} \times \overrightarrow{\Theta}, \mathscr{A} \times \overrightarrow{\mathscr{B}})$ 上的转移函数.

证明甚易. 读者可作为习题验证之.

**定理 1.1** 对于任何 $p - \Phi$ 链 $V = (\mathscr{X}, \mathscr{A}; \Theta, \mathscr{B}; p, \Phi)$, 存在一个概率空间 $(\Omega, \mathscr{F}, P_\Phi)$ 及定义在其上取值于 $\overrightarrow{\mathscr{X}} \times \overrightarrow{\Theta}$ 的 MCTRE $(\overrightarrow{X}, \overrightarrow{\xi})$, 使

$$P_\Phi\left((\overrightarrow{X}, \overrightarrow{\xi}) \in C\right) = \int_{\mathscr{X} \times \overrightarrow{\Theta}} \Phi(dx, d\overrightarrow{\theta}) \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x}) \mathbf{1}_C(\overrightarrow{x}, \overrightarrow{\theta}),$$
$$(C \in \overrightarrow{\mathscr{A}} \times \overrightarrow{\mathscr{B}}), \tag{1.9}$$

此外, 还有:

(1) $\overrightarrow{X}$ 的 $P_\Phi$ 边缘分布为

$$P_\Phi \circ (\overrightarrow{X})^{-1}(A) = \int_{\mathscr{X} \times \overrightarrow{\Theta}} \Phi(dx, d\overrightarrow{\theta}) P^x_{\overrightarrow{\theta}}(A), \quad (A \in \overrightarrow{\mathscr{A}}); \tag{1.10}$$

(2) $\overrightarrow{\xi}$ 的 $P_\Phi$ 边缘分布为

$$P_\Phi \circ (\overrightarrow{\xi})^{-1}(B) = \Phi(\mathscr{X} \times B) \overset{\text{def.}}{=\!=} \pi(B) \quad (B \in \overrightarrow{\mathscr{B}}); \tag{1.11}$$

(3) $\{(X_n, T^n \overrightarrow{\xi}), n \geqslant 0\}$ 是以 $\Phi$ 为初始分布, 以 $(\mathscr{X} \times \overrightarrow{\Theta}, \mathscr{A} \times \overrightarrow{\mathscr{B}})$ 为状态空间, 以 $\widehat{P}^{(n)}((x, \overrightarrow{\theta}), F)$ 为转移函数的时齐的 Markov 链, 称之为**斜积 Markov 链** (简记为 SPMC).

(4) 对任意 $(x, \overrightarrow{\theta}) \in \mathscr{X} \times \overrightarrow{\Theta}$, 定义 $\overrightarrow{\mathscr{A}} \times \overrightarrow{\mathscr{B}}$ 上的概率测度 $P_{(x, \overrightarrow{\theta})}$ 如下:

$$P_{(x, \overrightarrow{\theta})}(C) = \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x}) \mathbf{1}_C(\overrightarrow{x}, \overrightarrow{\theta}), \quad C \in \overrightarrow{\mathscr{A}} \times \overrightarrow{\mathscr{B}}, \tag{1.12}$$

则

$$P_\Phi((\vec{X},\vec{\xi})\in C)=P_\Phi(C)=\int_{\mathscr{X}\times\vec{\Theta}}\Phi(dx,d\vec{\theta})P_{(x,\vec{\theta})}(C),\tag{1.13}$$

$$P_{(x,\vec{\theta})}((X_0,\vec{\xi})=(x,\vec{\theta}))=1.\tag{1.14}$$

(5) $\vec{X}$ 是概率空间 $(\Omega,\mathscr{F},P_{(x,\vec{\theta})})$ 上的 $p-\vec{\theta}$ 链, 也是 $(\Omega,\mathscr{F},\overline{\mu}(\vec{\theta}))$ 上的 $p-\vec{\theta}$ 链, 其中

$$\overline{\mu}(\vec{\theta};C)=\int_{\mathscr{X}}\varphi(dx;\vec{\theta})\int_{\vec{\mathscr{X}}}P^x_{\vec{\theta}}(d\vec{x})\mathbf{1}_C(\vec{x},\vec{\theta}),$$
$$(\vec{\theta}\in\vec{\Theta};C\in\vec{\mathscr{A}}\times\vec{\mathscr{B}}),\tag{1.15}$$

$$\varphi(x,\vec{\theta})=P_\Phi(X_0=x|\vec{\xi}=\vec{\theta}),\tag{1.16}$$

此外, 对任何 $n\geqslant 1,x_n\in\mathscr{X},A_0,A_1,\cdots,A_{n-1}\in\mathscr{A},B\in\vec{\mathscr{A}}$, 有

$$\begin{aligned}&P_{(x,\vec{\theta})}((X_{n+1},X_{n+2},\cdots)\in B|X_i\in A_i,0\leqslant i<n,X_n=x_n)\\&=P_{(x_n,T^n\vec{\theta})}((X_1,X_2,\cdots)\in B).\end{aligned}\tag{1.17}$$

**证**　令 $\Omega=\vec{\mathscr{X}}\times\vec{\Theta},\mathscr{F}=\mathscr{A}\times\vec{\mathscr{B}}$,

$$P_\Phi(C)=\int_{\mathscr{X}\times\vec{\Theta}}\Phi(dx,d\vec{\theta})\int_{\vec{\mathscr{X}}}P^x_{\vec{\theta}}(d\vec{x})\mathbf{1}_C(\vec{x},\vec{\theta}),\quad (C\in\vec{\mathscr{A}}\times\vec{\mathscr{B}})\tag{1.18}$$

则 $(\Omega,\mathscr{F},P_\Phi)$ 是一个概率空间.

对任何 $\omega=(\vec{x},\vec{\theta})\in\Omega$, 定义

$$\vec{X}(\vec{x},\vec{\theta})=\vec{x},\quad \vec{\xi}(\vec{x},\vec{\theta})=\vec{\theta},\tag{1.19}$$

更详细地, 当 $\vec{x}=\{x_0,x_1,\cdots\},\vec{\theta}=\{\theta_n,n\in\mathbf{Z}\}$ 时, $X_k(\vec{x},\vec{\theta})=x_k,\xi_n(\vec{x},\vec{\theta})=\theta_n,(k\in\mathbf{N},n\in\mathbf{Z})$, 其中 $\vec{X}=\{X_0,X_1,\cdots\},\vec{\xi}=\{\xi_n,n\in\mathbf{Z}\}$.

现在证明 $(\vec{X},\vec{\xi})$ 是 $(\Omega,\mathscr{F},P_\Phi)$ 上的一个 MCRE. 令 $\Theta_n:\vec{\Theta}\mapsto\Theta$ 是第 $n$ 个坐标算子, 即 $\Theta_n(\vec{\theta})=\theta_n$ (当 $\vec{\theta}=(\theta_n,n\in\mathbf{Z})\in\vec{\Theta}$ 时). $\mathscr{B}^l_k$ 是 $\{\Theta_n,k-1<n<l+1\}$ 所产生的 $\sigma$ 代数 $(-\infty\leqslant k\leqslant l\leqslant\infty)$, 则

$$\sigma\{\xi_n,k-1<n<l+1\}=\{\vec{\mathscr{X}}\times B:B\in\mathscr{B}^l_k\}.\tag{1.20}$$

对任何 $B_i\in\mathscr{B},-\infty<k\leqslant l\leqslant n,n\geqslant 1,\{x_0,x_1,\cdots,x_{n+1}\}\subset\mathscr{X}$, 由 (1.18)、(1.20) 和 (1.7), 并注意 $X_n(\omega)=x_n,\xi_n(\omega)=\theta_n$ (当 $\omega=(\vec{x},\vec{\theta})$) 有:

$$\begin{aligned}&\int_{L(\vec{\mathscr{X}}:x_0,\cdots,x_n)\times L(\vec{\Theta}:B_k,\cdots,B_l)}\left[p(\xi_n;X_n,x_{n+1})\bullet P_\Phi\left(\xi_{n+1}\in B_{n+1}\left|\vec{\xi}^{\,n}_{-\infty}\right.\right)\right]dP_\Phi\\&=\int_{\mathscr{X}\times\vec{\Theta}}\Phi(dx,d\vec{\theta})\left[P^x_{\vec{\theta}}(L(\vec{x}:x_0,x_1,\cdots,x_{n+1}))\right.\\&\quad\left.\bullet\,\pi(L(\vec{\Theta}:B_k,\cdots,B_l,M_{l+1},\cdots,M_n,B_{n+1})|\mathscr{B}^n_{-\infty})\right].\end{aligned}\tag{1.21}$$

其中 $\pi(B)=\Phi(\mathscr{X}\times B)\ (B\in\mathscr{B}), M_{l+1}=\cdots=M_n=\Theta$.

但是, 对任何固定的 $x, x_0, x_1, \cdots, x_{n+1}\in\mathscr{X}$,

$$\begin{aligned}&P^{x}_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}}:x_0,x_1,\cdots,x_{n+1})\\&=\delta_{x,x_0}\prod_{i=0}^{n}p(\theta_i;x_i,x_{i+1})\end{aligned}$$

作为 $\overrightarrow{\theta}$ 的函数, 关于 $\mathscr{B}_0^n$ 可测, 而 $\mathscr{B}_0^n\subset\mathscr{B}_{-\infty}^n$, 所以 (1.21) 右端等于

$$\begin{aligned}&\int_{\mathscr{X}\times\overrightarrow{\Theta}}\Phi(dx,d\overrightarrow{\theta})\Big[\boldsymbol{E}_\pi\Big(P^{x}_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}}:x_0,x_1,\cdots,x_{n+1}))\\&\cdot\boldsymbol{1}_{L(\overrightarrow{\Theta}:B_k,\cdots,B_l,M_{l+1},\cdots,M_n,B_{n+1})}(\overrightarrow{\theta})\,\Big|\mathscr{B}_{-\infty}^n\Big)\Big]\\&=P_\Phi\left(\bigcap_{i=0}^{n+1}\{X_i=x_i\}\bigcap\bigcap_{j=k}^{l}\{\xi_j\in B_j\}\bigcap\{\xi_{n+1}\in B_{n+1}\}\right),\end{aligned}\qquad(1.22)$$

其中 $\boldsymbol{E}_\pi$ 是关于概率测度 $\pi$ 的期望算子. 由 (1.21) 和 (1.22) 得知 (1.5) 成立, 此即 $(\overrightarrow{X},\overrightarrow{\xi})$ 是一个以 $p$ 为随机转移矩阵的 MCTRE.

(1) 和 (2) 由 $P_\Phi, \overrightarrow{X}, \overrightarrow{\xi}$ 的定义 (见 (1.18) 和 (1.19)) 立即可得.

(3) 由 $P_\Phi$ 的定义

$$\begin{aligned}&P_\Phi\left(\bigcap_{i=0}^{n}\{X_i=x_i,T^i\overrightarrow{\xi}\in B_i\}\right)\\&=\int_{B_0}\Phi(x_0,d\overrightarrow{\theta})\left[P^{x_0}_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}}:x_0,\cdots,x_n))\prod_{i=1}^{n}\boldsymbol{1}_{B_i}(T^i\overrightarrow{\theta})\right]\\&\qquad(x_i\in\mathscr{X},B_i\in\mathscr{B},0\leqslant i\leqslant n),\end{aligned}\qquad(1.23)$$

所以为证 (3), 只需证明 (1.23) 右端方括号内的被积函数等于

$$\begin{aligned}&\int_{\mathscr{X}\times\overrightarrow{\Theta}}\widehat{P}((x_0,\overrightarrow{\theta}),d(y_1,\overrightarrow{\theta}_1))\int_{\mathscr{X}\times\overrightarrow{\Theta}}\widehat{P}((y_1,\overrightarrow{\theta}_1),d(y_2,\overrightarrow{\theta}_2))\cdots\\&\int_{\mathscr{X}\times\overrightarrow{\Theta}}\widehat{P}((y_{n-1},\overrightarrow{\theta}_{n-1}),d(y_n,\overrightarrow{\theta}_n))\prod_{i=1}^{n}\boldsymbol{1}_{B_i}(\overrightarrow{\theta}_i)\boldsymbol{1}_{\{x_i\}}(y_i).\end{aligned}\qquad(1.24)$$

由于转移函数 $\widehat{P}^{(n)}((x,\overrightarrow{\theta}),F)$ 的定义 (见 (1.8) 可知, 对 $(\mathscr{X}\times\overrightarrow{\Theta},\mathscr{A}\times\overrightarrow{\mathscr{B}})$ 上的任何有界可测函数 $f(x,\overrightarrow{\theta})$ 有

$$\begin{aligned}&\int_{\mathscr{X}\times\overrightarrow{\Theta}}\widehat{P}^{(n)}((x,\overrightarrow{\theta}),d(y,\overrightarrow{\tau}))f(y,\overrightarrow{\tau})\\&=\sum_{y\in\mathscr{X}}p^{(n)}(\overrightarrow{\theta};x,y)f(y,T^n\overrightarrow{\theta}).\end{aligned}\qquad(1.25)$$

因此, 由 (1.25) 和 (1.8) 知

$$
\begin{aligned}
&\int_{\mathscr{X}\times\vec{\Theta}}\widehat{P}((x_0,\vec{\theta}),d(y_1,\vec{\theta}_1))\int_{\mathscr{X}\times\vec{\Theta}}\widehat{P}((y_1,\vec{\theta}_1),d(y_2,\vec{\theta}_2))\\
&\cdot\prod_{i=1}^{2}\mathbf{1}_{B_i}(\vec{\theta}_i)\mathbf{1}_{\{x_i\}}(y_i).\\
=&\int_{\mathscr{X}\times\vec{\Theta}}\widehat{P}((x_0,\vec{\theta}),d(y_1,\vec{\theta}_1))\left[\widehat{P}((y_1,\vec{\theta}_1),\{x_2\}\times B_2)\cdot\mathbf{1}_{B_1}(\vec{\theta}_1)\mathbf{1}_{\{x_1\}}(y_1)\right]\\
=&\int_{\mathscr{X}\times\Theta}\widehat{P}((x_0,\vec{\theta}),d(y_1,\vec{\theta}_1))\Bigg[\mathbf{1}_{B_1}(\vec{\theta}_1)\mathbf{1}_{\{x_1\}}(y_1)\\
&\cdot\sum_{u\in\mathscr{X}}p^{(1)}(\vec{\theta}_1;y_1,u)\mathbf{1}_{(\{x_2\}\times B_2)_u}(T\vec{\theta}_1)\Bigg]\\
=&\int_{\mathscr{X}\times\vec{\Theta}}\widehat{P}((x_0,\vec{\theta}),d(y_1,\vec{\theta}_1))\Big[\mathbf{1}_{B_1}(\vec{\theta}_1)\mathbf{1}_{\{x_1\}}(y_1)\\
&\cdot p^{(1)}(\vec{\theta}_1;y_1,x_2)\mathbf{1}_{B_2}(T\vec{\theta}_1)\Big]\\
=&\sum_{y_1\in\mathscr{X}}p^{(1)}(\vec{\theta};x_0,y_1)[\mathbf{1}_{B_1}(T\vec{\theta})\mathbf{1}_{\{x_1\}}(y_1)\cdot p^{(1)}(T\vec{\theta};y_1,x_2)\mathbf{1}_{B_2}(T^2\vec{\theta})]\\
=&\,p^{(1)}(\vec{\theta};x_0,x_1)p^{(1)}(T\vec{\theta};x_1,x_2)\prod_{i=1}^{2}\mathbf{1}_{B_i}(T^i\vec{\theta})\\
=&\,p(\theta_0;x_0,x_1)p(\theta_1;x_1,x_2)\prod_{i=1}^{2}\mathbf{1}_{B_i}(T^i\vec{\theta})\\
=&\,P^{x_0}_{\vec{\theta}}(L(\vec{\mathscr{X}}:x_0,x_1,x_2))\prod_{i=1}^{2}\mathbf{1}_{B_i}(T^i\vec{\theta}).
\end{aligned}
$$

对 $n$ 作归纳法可证: (1.23) 右端方括号内之被积函数等于 (1.24) 式. 所以 (3) 获证.

(4) 是显然的.

(5) 由 (1.12) 有

$$
\begin{aligned}
&P_{(x,\vec{\theta})}(X_0=x_0,X_1=x_1,\cdots,X_n=x_n)\\
=&\,P^{x}_{\vec{\theta}}(L(\vec{\mathscr{X}}:x_0,x_1,\cdots,x_n))\\
=&\,\delta_{x,x_0}\prod_{i=0}^{n-1}p(\theta_i;x_i,x_{i+1})=\delta_{x,x_0}\prod_{i=0}^{n-1}p^{(1)}(T^i\vec{\theta};x_i,x_{i+1}),
\end{aligned}\tag{1.26}
$$

所以 $\vec{X}$ 是 $(\Omega,\mathscr{F},P_{(x,\vec{\theta})})$ 上的 $p-\vec{\theta}$ 链.

又因为对任何 $x_n, x_{n+1}, \cdots, x_{n+m} \in \mathscr{X}, A_i \in \mathscr{A}$ $(0 \leqslant i < n)$，有

$$\begin{aligned}
& P_{(x,\overrightarrow{\theta})}((X_{n+1}, X_{n+2}, \cdots) \in L(\overrightarrow{\mathscr{X}} : x_{n+1}, \cdots, x_{n+m}), \\
& X_n = x_n, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) \\
= \ & P_{(x,\overrightarrow{\theta})}(X_0 \in A_0, \cdots, X_{n-1} \in A_{n-1}, X_n = x_n, \cdots, X_{n+m} = x_{n+m}) \\
\overset{(1.12)}{=\!=} & \sum_{\substack{x_i \in A_i \\ 0 \leqslant i < n}} P^x_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}} : x_0, \cdots, x_{n-1}, x_n, \cdots, x_{n+m})) \\
= \ & \left[\sum_{\substack{x_i \in A_i \\ 0 \leqslant i < n}} \delta_{x,x_0} \prod_{i=0}^{n-1} p^{(1)}(T^i \overrightarrow{\theta}; x_i, x_{i+1})\right] \cdot \prod_{j=n}^{n+m-1} p^{(1)}(T^j \overrightarrow{\theta}; x_j, x_{j+1}) \\
= \ & \left[\sum_{\substack{x_i \in A_i \\ 0 \leqslant i < n}} \delta_{x,x_0} \prod_{i=0}^{n-1} p^{(1)}(T^i \overrightarrow{\theta}; x_i, x_{i+1})\right] \cdot \prod_{j=0}^{m-1} p^{(1)}(T^j (T^n \overrightarrow{\theta}); x_{n+j}, x_{n+j+1}) \\
= \ & \left[\sum_{\substack{x_i \in A_i \\ 0 \leqslant i < n}} \delta_{x,x_0} \prod_{i=0}^{n-1} p^{(1)}(T^i \overrightarrow{\theta}; x_i, x_{i+1})\right] \\
& \cdot P_{(x_n, T^n \overrightarrow{\theta})}((X_1, X_2, \cdots) \in L(\overrightarrow{\mathscr{X}} : x_{n+1}, \cdots, x_{n+m})). \qquad (1.27)
\end{aligned}$$

由 (1.26)、(1.27) 知

$$\begin{aligned}
& P_{(x,\overrightarrow{\theta})}((X_{n+1}, X_{n+2}, \cdots) \in L(\overrightarrow{\mathscr{X}} : x_{n+1}, \cdots, x_{n+m}) \\
& |X_n = x_n, X_{n-1} \in A_{n-1}, \cdots, X_0 \in A_0) \\
& = P_{(x_n, T^n \overrightarrow{\theta})}((X_1, X_2, \cdots) \in L(\overrightarrow{\mathscr{X}} : x_{n+1}, \cdots, x_{n+m})).
\end{aligned}$$

而 $\overrightarrow{\mathscr{A}}$ 是由一切柱集 $\{L(\overrightarrow{\mathscr{X}} : y_0, y_1, \cdots, y_n) : n \geqslant 0, y_i \in \mathscr{X}\}$ 所产生的 $\sigma$ 代数，所以，对任何 $n \geqslant 1, x_n \in \mathscr{X}, A_0, A_1, \cdots, A_{n-1} \in \mathscr{A}, B \in \overrightarrow{\mathscr{A}}$，都有

$$\begin{aligned}
& P_{(x,\overrightarrow{\theta})}((X_{n+1}, X_{n+2}, \cdots) \in B | X_i \in A_i, 0 \leqslant i < n, X_n = x_n) \\
& = P_{(x_n, T^n \overrightarrow{\theta})}((X_1, X_2, \cdots) \in B)
\end{aligned}$$

(只要左方有意义).

又因为由 (1.15)

$$\begin{aligned}
& \overline{\mu}(\overrightarrow{\theta})(X_0 = x_0, X_1 = x_1, \cdots, X_n = x_n) \\
& = \int_{\mathscr{X}} \varphi(dx, \overrightarrow{\theta}) P^x_{\overrightarrow{\theta}}(L(\overrightarrow{\mathscr{X}} : x_0, x_1, \cdots, x_n)) \\
& = \delta_{x,x_0} \prod_{i=0}^{n-1} p(\theta_i; x_i, x_{i+1}),
\end{aligned}$$

所以 $\overrightarrow{X}$ 也是 $(\Omega, \mathscr{F}, \overline{\mu}(\overrightarrow{\theta}))$ 上的 $p-\overrightarrow{\theta}$ 链. 定理证毕.

**定理 1.2**　$\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 是具有随机转移矩阵 $p$ 的 MCTRE 的充分必要条件是:

(1) $P(X_{n+1}=y|\overrightarrow{X}_0^n, \overrightarrow{\xi})=p(\xi_n; X_n, y)$;

(2) $P(X_0=x|\overrightarrow{\xi})=P(X_0=x|\overrightarrow{\xi}_{-\infty}^0)$, $(\forall n \geqslant 0, x \in \mathscr{X})$.

在 §6 中还要证明一个比这更一般的定理, 为避免重复, 定理 1.2 暂不证. (参看定理 6.4).

**定义 1.8**　称定理 1.1 中构造出的 $(\overrightarrow{X}, \overrightarrow{\xi})$ 为由 $p-\Phi$ 链 $V=(\mathscr{X}, \mathscr{A}; \Theta, \mathscr{B}; p, \Phi)$ 所产生的随机环境中的 Markov 链 (MCRE), 而称 $\{(X_n, T^n\overrightarrow{\xi}), n \geqslant 0\}$ 为由 $V$ 产生的斜积 Markov 链 (SPMC).

**附注 1.2**　一般地 $\{X_n, n \geqslant 0\}$ 未必是 $(\Omega, \mathscr{F}, P_\Phi)$ 上的 Markov 链, 但是, 它是概率空间 $(\Omega, \mathscr{F}, \overline{\mu}(\overrightarrow{\theta}))$ 上的 $p-\overrightarrow{\theta}$ 链, 是非时齐的 Markov 链. 而 SPMC $\{(X_n, T^n\overrightarrow{\xi}), n \geqslant 0\}$ 是一般状态 (未必可数) 的时齐的 Markov 链. 这一事实很重要, 它可以借用来研究 MCTRE. MCRE $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 既可视为两个随机过程, 也可以耦合在一起看成一个取值于 $\overrightarrow{\mathscr{X}} \times \overrightarrow{\Theta}$ 的随机元.

**附注 1.3**　定理 1.1 中所构造的 MCTRE $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 中的随机环境 $\overrightarrow{\xi}=\{\xi_n, n \in \mathbf{Z}\}$ 仅依赖 "时间" $n$, 并不依赖 $X_n$ 的空间位置, 故称 $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 是依时随机环境中的 Markov 链.

## §2　依时随机环境中的 Markov 链的特性函数及其性质

本节沿袭 §1 中的符号. $V=(\mathscr{X}, \mathscr{A}; \Theta, \mathscr{B}; p, \Phi)$ 是一个 $p-\Phi$ 链, $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 是定理 1.1 中所构造的 MCTRE, $\{(X_n, T^n\overrightarrow{\xi}), n \geqslant 0\}$ 是对应的 SPMC.

$(\Omega, \mathscr{F}, P_\Phi), T, P_{\overrightarrow{\theta}}^x, P_{(x,\overrightarrow{\theta})}, \widehat{P}^{(n)}((x, \overrightarrow{\theta}), F)$ 如 §1 所定义. 记 $\eta_n=(X_n, T^n, \overrightarrow{\xi})$ $(n \geqslant 0)$.

$$(F)_y=\{\overrightarrow{\theta} \in \overrightarrow{\Theta}: (y, \overrightarrow{\theta}) \in F\}, \quad F \in \mathscr{A} \times \overrightarrow{\mathscr{B}},$$
$$(F)^{\overrightarrow{\theta}}=\{y \in \mathscr{X}: (y, \overrightarrow{\theta}) \in F\}, \quad F \in \mathscr{A} \times \overrightarrow{\mathscr{B}}$$

是截口集.

$p^{(n)}(\overrightarrow{\theta}; x, y)$　如 (1.2) 式所定义.

**命题 2.1**　对任何 $(x, \overrightarrow{\theta}) \in \mathscr{X} \times \overrightarrow{\Theta}, n \geqslant 1, G \in \mathscr{A} \times \overrightarrow{\mathscr{B}}$, 恒有

$$\widehat{P}^{(n)}((x, \overrightarrow{\theta}), G)=P_{(x,\overrightarrow{\theta})}(\eta_n \in G). \tag{2.1}$$

**证** 由于 $\mathscr{X}$ 是可数集, 为证 (2.1), 对一切 $G \in \mathscr{A} \times \overrightarrow{\mathscr{B}}$ 成立, 只需证 (2.1) 对 $G = \{y\} \times B(y \in \mathscr{X}, B \in \overrightarrow{\mathscr{B}})$ 成立即可. 事实上, 由 (1.8)* 及 (1.12) 知:

$$
\begin{aligned}
&P_{(x,\overrightarrow{\theta})}(\eta_n \in \{y\} \times B) \\
\overset{(1.12)}{=\!=\!=}\ & \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x})\mathbf{1}_{\{\eta_n \in \{y\}\times B\}}(\overrightarrow{x}, \overrightarrow{\theta}) \\
=\ & \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x})\mathbf{1}_{\{X_n = y, T^n\overrightarrow{\xi} \in B\}}(\overrightarrow{x}, \overrightarrow{\theta}) \\
=\ & \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x})\mathbf{1}_{\{\overrightarrow{x} \in \overrightarrow{\mathscr{X}}: x_n = y\}}(\overrightarrow{x})\mathbf{1}_B(T^n\overrightarrow{\theta}) \\
=\ & p^{(n)}(\overrightarrow{\theta}; x, y)\mathbf{1}_B(T^n\overrightarrow{\theta}) \\
\overset{(1.8)^*}{=\!=\!=}\ & \widehat{P}^{(n)}((x, \overrightarrow{\theta}), \{y\} \times B).
\end{aligned}
\tag{2.2}
$$

命题 2.1 得证.

由 (2.2) 立得

$$
\begin{aligned}
\widehat{P}^{(n)}((x, \overrightarrow{\theta}), \{y\} \times \overrightarrow{\Theta}) &= p^{(n)}(\overrightarrow{\theta}; x, y) \\
&= P_{(x,\overrightarrow{\theta})}(X_n = y).
\end{aligned}
\tag{2.3}
$$

类似于经典 (固定环境) 的时齐的 Markov 链, 我们可以令

$$
\widehat{F}^{(n)}((x, \overrightarrow{\theta}), G) = P_{(x,\overrightarrow{\theta})} \quad (\eta_n \in G, \eta_k \overline{\in} G, 1 \leqslant k < n) \tag{2.4}
$$

为 $\{\eta_n, n \geqslant 0\}$ 从 $(x, \overrightarrow{\theta})$ 出发在时刻 $n$ 初达 $G$ 的概率.

(注意: $\widehat{P}^{(n)}((x, \overrightarrow{\theta}), G)$ 是 $\{\eta_n, n \geqslant 0\}$ 从 $(x, \overrightarrow{\theta})$ 出发在时刻 $n$ 处于状态集合 $G$ 的概率.) 令

$$
\begin{aligned}
f^{(n)}(\overrightarrow{\theta}; x, y) &\overset{\text{def.}}{=\!=} \widehat{F}^{(n)}((x, \overrightarrow{\theta}), \{y\} \times \overrightarrow{\Theta}) \\
&= P_{(x,\overrightarrow{\theta})}(X_n = y, X_k \neq y, 1 \leqslant k < n)
\end{aligned}
\tag{2.5}
$$

为在环境 $\overrightarrow{\xi} = \overrightarrow{\theta}$ 的条件下, $\{X_n, n \geqslant 0\}$ 从 $x$ 出发, 在时刻 $n$ 初到达状态 $y$ 的概率.

(注意: $p^{(n)}(\overrightarrow{\theta}; x, y)$ 是 $\{X_n, n \geqslant 0\}$ 在环境 $\overrightarrow{\xi} = \overrightarrow{\theta}$ 的条件下, $\{X_n, n \geqslant 0\}$ 从 $x$ 出发, 在时刻 $n$ 处于状态 $y$ 的概率.)

再令

$$
\widehat{F}^*((x, \overrightarrow{\theta}), G) = \sum_{n=1}^{\infty} \widehat{F}^{(n)}((x, \overrightarrow{\theta}), G), \tag{2.6}
$$

$$
f^*(\overrightarrow{\theta}; x, y) = \sum_{n=1}^{\infty} f^{(n)}(\overrightarrow{\theta}; x, y), \tag{2.7}
$$

注意: $\widehat{F}^*((x,\overrightarrow{\theta}),G)$ 是 $\{\eta_n, n\geqslant 0\}$ 从 $(x,\overrightarrow{\theta})$ 出发, 经有限步到达 $G$ 的概率; $f^*(\overrightarrow{\theta};x,y)$ 是 $\{X_n, n\geqslant 0\}$ 在环境 $\overrightarrow{\xi}=\overrightarrow{\theta}$ 的条件下, 从 $x$ 出发, 经有限步到达状态 $y$ 的概率. 再令

$$Q((x,\overrightarrow{\theta}),G)=P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{\eta_k\in G\}\right), \tag{2.8}$$

是 $\{\eta_n, n\geqslant 0\}$ 从 $(x,\overrightarrow{\theta})$ 出发无穷多次进入 $G$ 的概率 $((x,\overrightarrow{\theta})\in\mathscr{X}\times\overrightarrow{\Theta}, G\in\mathscr{A}\times\overrightarrow{\mathscr{B}})$,

$$\begin{aligned}q(\overrightarrow{\theta};x,y)&\xlongequal{\text{def.}}Q((x,\overrightarrow{\theta}),\{y\}\times\overrightarrow{\Theta})\\&=P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=y\}\right)\end{aligned} \tag{2.9}$$

是 $\{X_n, n\geqslant 0\}$ 在环境 $\overrightarrow{\xi}=\overrightarrow{\theta}$ 的条件下, 从 $x$ 出发无穷多次地进入状态 $y$ 的概率 $(x,y\in\mathscr{X},\overrightarrow{\theta}\in\overrightarrow{\Theta})$.

注意: $P^x_{\overrightarrow{\theta}}(A)$ 对任何固定的 $x\in\mathscr{X}, A\in\overrightarrow{\mathscr{A}}$ 是 $\overrightarrow{\theta}$ 的关于 $\sigma$ 代数 $\overrightarrow{\mathscr{B}}$ 的可测函数, 所以, 由 $P_{(x,\overrightarrow{\theta})}(C)$ 的定义 (见 (1.12)) 可知: 对任何固定的 $x\in\mathscr{X}, C\in\overrightarrow{\mathscr{A}}\times\overrightarrow{\mathscr{B}}$, 它是 $\overrightarrow{\theta}$ 的关于 $\sigma$ 代数 $\overrightarrow{\mathscr{B}}$ 的可测函数. 如不声明, 本节一律使用上述符号.

**命题 2.2** 我们恒有:

$$q(\overrightarrow{\theta};x,y)=\sum_{r=1}^{\infty}f^{(r)}(\overrightarrow{\theta};x,y)q(T^r\overrightarrow{\theta};y,y)\quad(\overrightarrow{\theta}\in\overrightarrow{\Theta};x,y\in\mathscr{X}). \tag{2.10}$$

**证** 由 (1.17) 有

$$\begin{aligned}q(\overrightarrow{\theta};x,y)&=P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=y\}\right)\\
&=\sum_{r=1}^{\infty}P_{(x,\overrightarrow{\theta})}\left(X_k\neq y,1\leqslant k<r,X_r=y,\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\{X_{r+m}=y\}\right)\\
&=\sum_{r=1}^{n}P_{(x,\overrightarrow{\theta})}(X_k\neq y,1\leqslant k<r,X_r=y)\\
&\quad\cdot P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\{X_{r+m}=y\}\,|X_k\neq y,1\leqslant k<r,X_r=y\right)\\
&\xlongequal{(1.17)}\sum_{r=1}^{\infty}f^{(r)}(\overrightarrow{\theta};x,y)P_{(y,T^r\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\{X_m=y\}\right)\\
&=\sum_{r=1}^{\infty}f^{(r)}(\overrightarrow{\theta};x,y)q(T^r\overrightarrow{\theta};y,y).\end{aligned}$$

**命题 2.3** 对任何 $n \geqslant 1, x, y \in \mathscr{X}, \vec{\theta} \in \vec{\Theta}$, 恒有

$$p^{(n)}(\vec{\theta};x,y) = \sum_{k=1}^{n} f^{(k)}(\vec{\theta};x,y)p^{(n-k)}(T^k\vec{\theta};y,y)$$
$$= f^{(n)}(\vec{\theta};x,y) + \sum_{k=1}^{n-1} p^{(k)}(\vec{\theta};x,y)f^{(n-k)}(T^k\vec{\theta};y,y). \tag{2.11}$$

**证** 由 $p^{(n)}(\vec{\theta};x,y) = \widehat{P}^{(n)}((x,\vec{\theta}),\{y\}\times\vec{\Theta})$ 及 (2.1) 和 (1.17) 知

$$\begin{aligned}
p^{(n)}(\vec{\theta};x,y) &= P_{(x,\vec{\theta})}(X_n = y) \\
&= \sum_{k=1}^{n} P_{(x,\vec{\theta})}(X_s \neq y, 1 \leqslant s < k, X_k = y) \\
&\quad \bullet P_{(x,\vec{\theta})}(X_n = y | X_s \neq y, 1 \leqslant s < k, X_k = y) \\
&\overset{(1.17)}{=\!=} \sum_{k=1}^{n} P_{(x,\vec{\theta})}(X_s \neq y, 1 \leqslant s < k, X_k = y) \\
&\quad \bullet P_{(y,T^k\vec{\theta})}(X_{n-k} = y) \\
&= \sum_{k=1}^{n} f^{(k)}(\vec{\theta};x,y)p^{(n-k)}(T^k\vec{\theta};y,y).
\end{aligned} \tag{2.12}$$

由 (2.12) 得知 (2.11) 的第一个等式成立. 仿之可证 (2.11) 的第二个等式亦成立. 命题证毕.

再引一对 Green 函数和一对母函数. 令

$$G(\vec{\theta};x,F) \overset{\text{def.}}{=\!=} \sum_{n=0}^{\infty} P_{(x,\vec{\theta})}(\eta_n \in F)$$
$$= \sum_{n=0}^{\infty} \widehat{P}^{(n)}((x,\vec{\theta}),F) \ (x \in \mathscr{X}, \vec{\theta} \in \vec{\Theta}, F \in \mathscr{A} \times \mathscr{B}) \tag{2.13}$$

$$g(\vec{\theta};x,y) \overset{\text{def.}}{=\!=} G(\vec{\theta};x,\{y\}\times\vec{\Theta})$$
$$= \sum_{n=0}^{\infty} p^{(n)}(\vec{\theta};x,y), (x,y \in \mathscr{X}, \vec{\theta} \in \vec{\Theta}); \tag{2.14}$$

$$m_p(\vec{\theta};x,y,\lambda) \overset{\text{def.}}{=\!=} \sum_{n=0}^{\infty} p^{(n)}(\vec{\theta};x,y)\lambda^n, (\vec{\theta} \in \vec{\Theta}, x,y \in \mathscr{X}, |\lambda| \leqslant 1), \tag{2.15}$$

$$m_f(\vec{\theta};x,y,\lambda) \overset{\text{def.}}{=\!=} \sum_{n=1}^{\infty} f^{(n)}(\vec{\theta};x,y)\lambda^n, (\vec{\theta} \in \vec{\Theta}, x,y \in \mathscr{X}, |\lambda| \leqslant 1). \tag{2.16}$$

**命题 2.4**　对 (2.15)、(2.16) 定义的母函数, 恒有

$$(1)\quad m_p(\overrightarrow{\theta};x,y,\lambda)=\delta_{x,y}+\sum_{k=1}^{\infty}f^{(k)}(\overrightarrow{\theta};x,y)\lambda^k m_p(T^k\overrightarrow{\theta};y,y,\lambda); \tag{2.17}$$

$$(2)\quad \lim_{\lambda\to 1-0}m_f(\overrightarrow{\theta};x,y,\lambda)=m_f(\overrightarrow{\theta};x,y,1)=f^*(\overrightarrow{\theta};x,y); \tag{2.18}$$

$$(3)\quad \lim_{\lambda\to 1-0}m_p(\overrightarrow{\theta};x,y,\lambda)=\sum_{n=0}^{\infty}p^{(n)}(\overrightarrow{\theta};x,y)=g(\overrightarrow{\theta};x,y). \tag{2.19}$$

**证**　由命题 2.4 由命题 2.3 及 Green 函数和母函数的定义立得.

令

$$f^*_{\min}(\overrightarrow{\theta};x,y)=\inf_{k\in\mathbf{Z}}f^*(T^k\overrightarrow{\theta};x,y); \tag{2.20}$$

$$q_{\min}(\overrightarrow{\theta};x,y)=\inf_{k\in\mathbf{Z}}q(T^k\overrightarrow{\theta};x,y); \tag{2.21}$$

$$g_{\min}(\overrightarrow{\theta};x,y)=\inf_{k\in\mathbf{Z}}g(T^k\overrightarrow{\theta};x,y). \tag{2.22}$$

**命题 2.5**　若 $f^*_{\min}(\overrightarrow{\theta};x,y)>0$, 则 $q(\overrightarrow{\theta};x,y)\geqslant q(\overrightarrow{\theta};x,x)$ $(\forall\overrightarrow{\theta}\in\overrightarrow{\Theta},x,y\in\mathscr{X})$.

**证**　令

$$l_{(x,\overrightarrow{\theta})}(y,s,n)=P_{(x,\overrightarrow{\theta})}\left(\bigcup_{t=s}^{\infty}\{X_t=x\},\bigcap_{k=n}^{\infty}\{X_k\neq y\}\right)$$
$$(x,y\in\mathscr{X},\overrightarrow{\theta}\in\overrightarrow{\Theta},s>n),$$

则

$$\begin{aligned}q(\overrightarrow{\theta};x,y)&=P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=y\}\right)\\&\geqslant P_{(x,\overrightarrow{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=y\},\bigcap_{s=1}^{\infty}\bigcup_{t=s}^{\infty}\{X_t=x\}\right)\\&=P_{(x,\overrightarrow{\theta})}\left(\bigcap_{s=1}^{\infty}\bigcup_{t=s}^{\infty}\{X_t=x\}\right)\\&\quad-P_{(x,\overrightarrow{\theta})}\left(\bigcap_{s=1}^{\infty}\bigcup_{t=s}^{\infty}\{X_t=x\},\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{X_k\neq y\}\right)\\&=q(\overrightarrow{\theta};x,x)-\lim_{n\to\infty}\lim_{s\to\infty}l_{(x,\overrightarrow{\theta})}(y,s,n)\end{aligned} \tag{2.23}$$

因为 $\{X_n, n \geqslant 0\}$ 是 $(\Omega, \mathscr{F}, P_{(x,\overrightarrow{\theta})})$ 上的 $p-\overrightarrow{\theta}$ 链, 由 (1.17) 知:

$$\begin{aligned}
&l_{(x,\overrightarrow{\theta})}(y,s,n)\\
\leqslant\ & \sum_{t=s}^{\infty} P_{(x,\overrightarrow{\theta})}\left(X_u \neq x, s \leqslant u < t, X_t = x, \bigcap_{k=n}^{\infty}\{X_k \neq y\}\right)\\
\leqslant\ & \sum_{t=s}^{\infty} P_{(x,\overrightarrow{\theta})}\Bigg(X_v \neq y, n \leqslant v < s, X_u \neq x, s \leqslant u < t, X_t = x,\\
& \bigcap_{k=t+1}^{\infty}\{X_k \neq y\}\Bigg)\\
\xlongequal{(1.17)}\ & \sum_{t=s}^{\infty} P_{(x,\overrightarrow{\theta})}(X_v \neq y, n \leqslant v < s, X_u \neq x, s \leqslant u < t, X_t = x)\\
& \cdot P_{(x,T^t\overrightarrow{\theta})}\left(\bigcap_{j=1}^{\infty}\{X_j \neq y\}\right)\\
=\ & \sum_{t=s}^{\infty} P_{(x,\overrightarrow{\theta})}\left(\bigcap_{v=n}^{s-1}\{X_v \neq y\}, \bigcap_{u=s}^{t-1}\{X_u \neq x\}, X_t = x\right)\\
& \cdot(1 - f^*(T^t\overrightarrow{\theta}; x, y)\\
\leqslant\ & (1 - f^*_{\min}(\overrightarrow{\theta}; x, y))\\
& \cdot \sum_{t=s}^{\infty} P_{(x,\overrightarrow{\theta})}\left(\bigcap_{v=n}^{s-1}\{X_v \neq y\}, \bigcap_{u=s}^{t-1}\{X_u \neq x\}\{X_t = x\}\right)\\
\leqslant\ & (1 - f^*_{\min}(\overrightarrow{\theta}; x, y))P_{(x,\overrightarrow{\theta})}\left(\bigcap_{v=n}^{s-1}\{X_v \neq y\}, \bigcup_{t=s}^{\infty}\{X_t = x\}\right)\\
\leqslant\ & (1 - f^*_{\min}(\overrightarrow{\theta}; x, y))l_{(x,\overrightarrow{\theta})}(y,s,n).
\end{aligned} \tag{2.24}$$

在 (2.24) 中先令 $s \to \infty$ 再令 $n \to \infty$ 得

$$\lim_{n\to\infty}\lim_{s\to\infty} l_{(x,\overrightarrow{\theta})}(y,s,n) \leqslant (1 - f^*_{\min}(\overrightarrow{\theta}; x, y)).$$

但是

$$(1 - f^*_{\min}(\overrightarrow{\theta}; x, y)) > 0,$$

所以

$$\lim_{n\to\infty}\lim_{s\to\infty} l_{(x,\overrightarrow{\theta})}(y,s,n) = 0, \tag{2.25}$$

由 (2.24) 和 (2.25) 即得命题 2.5.

**命题 2.6** 若 $f^*_{\min}(\overrightarrow{\theta}; x, x) = 1$, 则 $q(\overrightarrow{\theta}; y, x) = f^*(\overrightarrow{\theta}; y, x)$ $(\forall y \in \mathscr{X})$ 且 $q_{\min}(\overrightarrow{\theta}; x, x) = 1$.

**证** 设 $f^*_{\min}(\overrightarrow{\theta};x,x)=1$. 令

$$\tau_x(1)=\inf\{n\geqslant 1: X_u\neq x, 1\leqslant u<n, X_n=x\},$$
$$\tau_x(k)=\inf\{n>\tau_x(k-1): X_u\neq x, 1\leqslant u<n, X_n=x\}\quad (k\geqslant 2),$$
$$\Gamma_x=\#\{n\geqslant 1: X_n=x\},$$

则

$$P_{(x,T^s\overrightarrow{\theta})}(\tau_x(1)<\infty)=f^*(T^s\overrightarrow{\theta};x,x)=1\quad (\forall s\geqslant 0).$$

若

$$P_{(x,T^s\overrightarrow{\theta})}(\tau_x(k)<\infty)=1\quad (\forall s\geqslant 0),$$

则由 (1.17) 可得

$$\begin{aligned}P_{(x,T^s\overrightarrow{\theta})}(\tau_x(k+1)<\infty)&=\sum_{n=1}^{\infty}f^{(n)}(T^s\overrightarrow{\theta};x,x)P_{(x,T^{s+n})}(\tau_x(k)<\infty)\\&=\sum_{n=1}^{\infty}f^{(n)}(T^s\overrightarrow{\theta};x,x)=f^*(T^s\theta;x,x)=1\quad(\forall s\geqslant 0).\end{aligned}$$

所以对一切 $k\geqslant 1$, 都有

$$P_{(x,T^s\overrightarrow{\theta})}(\tau_x(k))=1,$$

从而

$$\begin{aligned}q(T^n\overrightarrow{\theta};x,x)&=\lim_{k\to\infty}P_{(x,T^n\overrightarrow{\theta})}(\Gamma_x\geqslant k)\\&=\lim_{k\to\infty}P_{(x,T^n\overrightarrow{\theta})}(\tau_x(k)<\infty)=1\quad(\forall n\geqslant 0),\end{aligned}$$

所以 $q_{\min}(\overrightarrow{\theta};x,x)=1$. 显然由命题 2.2 还有

$$\begin{aligned}q(\overrightarrow{\theta};y,x)&=\sum_{n=1}^{\infty}f^{(n)}(\overrightarrow{\theta};y,x)q(T^n\overrightarrow{\theta};x,x)\\&=\sum_{n=1}^{\infty}f^{(n)}(\overrightarrow{\theta};y,x)=f^*(\overrightarrow{\theta};y,x).\end{aligned}$$

**命题 2.7** 若 $f^*_{\min}(\overrightarrow{\theta};x,x)=1$ 且 $f^*_{\min}(\theta^*;x,y)>0$, 则

$$f^*_{\min}(\overrightarrow{\theta};x,y)=q_{\min}(\overrightarrow{\theta};x,y)=q_{\min}(\overrightarrow{\theta};x,x)=1.$$

**证** 显然, $f^*_{\min}(\overrightarrow{\theta};x,y)\geqslant q_{\min}(\overrightarrow{\theta};x,y)$. 又因为 $f^*_{\min}(\overrightarrow{\theta};x,y)>0$, 所以由命题 2.5 有

$$q_{\min}(\overrightarrow{\theta};x,y)\geqslant q_{\min}(\overrightarrow{\theta};x,x).\tag{2.26}$$

而今假设 $f^*_{\min}(\overrightarrow{\theta};x,x)=1$, 则命题 2.6 知

$$q_{\min}(\overrightarrow{\theta};x,x)=1. \tag{2.27}$$

由 $f^*_{\min}(\overrightarrow{\theta};x,y)\geqslant q_{\min}(\overrightarrow{\theta};x,y)$ 和 (2.26)、(2.27) 知命题 2.7 成立.

**命题 2.8** 对任何 $\overrightarrow{\theta}\in\overrightarrow{\Theta},x,y\in\mathscr{X}$, 恒有

$$f^*(\overrightarrow{\theta};x,y)=p^{(1)}(\overrightarrow{\theta};x,y)+\sum_{u\neq y}p^{(1)}(\overrightarrow{\theta};x,u)f^*(T\overrightarrow{\theta};u,y). \tag{2.28}$$

**证** 由 $f^*(\overrightarrow{\theta};x,y)$、$p^{(1)}(\overrightarrow{\theta};x,y)$ 的定义和 (1.17) 式立即可证命题 2.8 成立.

**命题 2.9** 对任何 $\overrightarrow{\theta}\in\overrightarrow{\Theta},x,y\in\mathscr{X},C\in\mathscr{A}\times\overrightarrow{\mathscr{B}}$, 总有

$$\begin{aligned}Q(\overrightarrow{\theta};x,C)&=\widehat{F}^*(\overrightarrow{\theta};x,C)\\&-\sum_{n=1}^{\infty}\sum_{y\in\mathscr{X}}p^{(n)}(\overrightarrow{\theta};x,y)\bullet\mathbf{1}_{(C)_y}(T^n\overrightarrow{\theta})(1-\widehat{F}^*(T^n\overrightarrow{\theta},y,C)).\end{aligned} \tag{2.29}$$

**证** 记 $\boldsymbol{E}_{(x,\overrightarrow{\theta})}$ 为关于 $P_{(x,\overrightarrow{\theta})}$ 的期望算子, 则 (2.29) 左方为:

$$\begin{aligned}
&= 1-P_{(x,\overrightarrow{\theta})}\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{(X_k,T^k\overrightarrow{\theta})\overline{\in}C\}\right)\\
&= 1-P_{(x,\overrightarrow{\theta})}\left(\bigcap_{k=1}^{\infty}\{(X_k,T^k\overrightarrow{\theta})\overline{\in}C\}\right)\\
&\quad -\sum_{n=1}^{\infty}P_{(x,\overrightarrow{\theta})}((X_n,T^n\overrightarrow{\theta})\in C,(X_k,T^k\overrightarrow{\theta})\overline{\in}C,k>n)\\
&= \widehat{F}^*(\overrightarrow{\theta};x,C)-\sum_{n=1}^{\infty}\boldsymbol{E}_{(x,\overrightarrow{\theta})}\Big[\mathbf{1}_{\{(X_n,T^n\overrightarrow{\theta})\in C\}}\\
&\quad \bullet P_{(x,\overrightarrow{\theta})}((X_k,T^k\overrightarrow{\theta})\overline{\in}C,k>n|(X_n,T^n\overrightarrow{\theta}))\Big]\\
&= \widehat{F}^*(\overrightarrow{\theta};X,C)-\sum_{n=1}^{\infty}\boldsymbol{E}_{(x,\overrightarrow{\theta})}\left[\mathbf{1}_{\{(X_n,T^n\overrightarrow{\theta})\in C\}}\bullet(1-\widehat{F}^*(T^n\overrightarrow{\theta};X_n,C))\right]\\
&= \widehat{F}^*(\overrightarrow{\theta};x,C)-\sum_{n=1}^{\infty}\sum_{y\in\mathscr{X}}\boldsymbol{E}_{(x,\overrightarrow{\theta})}\left(\mathbf{1}_{\{y\}}(X_n)\mathbf{1}_{(C)_y}(T^n\overrightarrow{\theta})\right)\\
&\quad \bullet(1-\widehat{F}^*(T^n\overrightarrow{\theta};y,C))\\
&\overset{(2.3)}{=} \widehat{F}^*(\overrightarrow{\theta};x,C)-\sum_{n=1}^{\infty}\sum_{y\in\mathscr{X}}p^{(n)}(\overrightarrow{\theta};x,y)\mathbf{1}_{(C)_y}(T^n\overrightarrow{\theta})\\
&\quad \bullet(1-\widehat{F}^*(T^n\overrightarrow{\theta};y,C)).
\end{aligned}$$

(2.29) 证毕.

在 (2.29) 中取 $C=\{y\}\times\overrightarrow{\Theta}$ 得

$$q(\overrightarrow{\theta};x,y)=f^*(\overrightarrow{\theta};x,y)-\sum_{n=1}^{\infty}p^{(n)}(\overrightarrow{\theta};x,y)(1-f^*(T^n\overrightarrow{\theta};y,y)). \tag{2.29)*}$$

**命题 2.10** 对任何 $x\in\mathscr{X},\overrightarrow{\theta}\in\overrightarrow{\Theta},B\in\overrightarrow{\mathscr{B}}$, 有

$$\widehat{F}^*(\overrightarrow{\theta};x,\mathscr{X}\times\mathscr{B})=\mathbf{1}_{\bigcup\limits_{n=1}^{\infty}T^{-n}B}(\overrightarrow{\theta}). \tag{2.30}$$

**证** 由于 $P_{(x,\overrightarrow{\theta})}(\overrightarrow{\xi}=\overrightarrow{\theta})=1$, 所以

$$P_{(x,\overrightarrow{\theta})}(\overrightarrow{\xi}\in C)=\mathbf{1}_C(\overrightarrow{\theta})\quad(\forall C\in\overrightarrow{\mathscr{B}}).$$

因此

$$\begin{aligned}
&\widehat{F}^*(\overrightarrow{\theta};x,\mathscr{X}\times\mathscr{B})\\
&=P_{(x,\overrightarrow{\theta})}\left(\bigcup_{n=1}^{\infty}\{(X_n,T^n\overrightarrow{\xi})\in\mathscr{X}\times\mathscr{B}\}\right)\\
&=P_{(x,\overrightarrow{\theta})}\left(\bigcup_{n=1}^{\infty}\{T^n\overrightarrow{\xi}\in B\}\right)\\
&=P_{(x,\overrightarrow{\theta})}\left(\bigcup_{n=1}^{\infty}\{\overrightarrow{\xi}\in T^{-n}B\}\right)\\
&=P_{(x,\overrightarrow{\theta})}\left(\overrightarrow{\xi}\in\bigcup_{n=1}^{\infty}T^{-n}B\right)\\
&=\mathbf{1}_{\bigcup\limits_{n=1}^{\infty}T^{-n}B}(\overrightarrow{\theta}).
\end{aligned}$$

**命题 2.11** 设 $\pi(B)=\Phi(\mathscr{X}\times\mathscr{B}),B\in\overrightarrow{\mathscr{B}}$, 即 $\pi$ 是 $\overrightarrow{\xi}$ 的 $P_\Phi$ 边缘分布. 若 $\pi=\pi\circ T^{-1}$ (注意 $\pi$ 是 $(\overrightarrow{\Theta},\overrightarrow{\mathscr{B}})$ 上的概率测度, $T$ 是 $\overrightarrow{\Theta}$ 到 $\overrightarrow{\Theta}$ 的推移算子, 当然是 $(\overrightarrow{\Theta},\overrightarrow{\mathscr{B}},\pi)$ 上的随机元, $\pi\circ T^{-1}$ 就是 $T$ 的 $\pi$ 分布), 则下列陈述等价:

(1) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:f^*(\overrightarrow{\theta};x,x)=1\})=1$;

(2) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:q(\overrightarrow{\theta};x,x)=1\})=1$;

(3) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:f^*_{\min}(\overrightarrow{\theta};x,x)=1\})=1$;

(4) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:q_{\min}(\overrightarrow{\theta};x,x)=1\})=1$.

**证** 先证: (1) $\Rightarrow$ (2) 和 (3). 设 (1) 成立, 则可取 $A\in\overrightarrow{\mathscr{B}},\pi(A)=0$ 使

$$f^*(\overrightarrow{\theta};x,x)=1\quad(\forall\overrightarrow{\theta}\overline{\in}A).$$

令 $B=\bigcup\limits_{n\in\mathbf{Z}}T^{-n}A$, 则 $\pi(B)=0$, 且由 (2.29)* 和 (1) 成立有

$$q(\overrightarrow{\theta};x,x)=1\quad(\forall\overrightarrow{\theta}\overline{\in}B),$$

此即 (2) 成立. 又

$$\begin{aligned}&\pi(\{\overrightarrow{\theta}:f^{*}_{\min}(\overrightarrow{\theta};x,x)=1\})\\&=\pi\left(\bigcap_{n\in\mathbf{Z}}\{f^{*}(T^{n}\overrightarrow{\theta};x,x)=1\}=1,\right.\end{aligned}$$

此即 (3) 成立.

因为

$$f^{*}(\overrightarrow{\theta};x,x)\geqslant\max\{q(\overrightarrow{\theta};x,x),f^{*}_{\min}(\overrightarrow{\theta};x,x)\},$$

故由命题 2.6 可证

$$(2)\Rightarrow(1);\quad(3)\Rightarrow(1);\quad(3)\Rightarrow(4).$$

而 “(4) ⇒ (3)” 是显然的. 命题证毕.

**命题 2.12** 若 $\pi=\pi\circ T^{-1},\pi(\{\overrightarrow{\theta}:f^{*}(\overrightarrow{\theta};x,y)>0\})=1$, $\pi(\{\overrightarrow{\theta}:f^{*}(\overrightarrow{\theta};x,x)=1\})=1$, 则

$$\pi(\{\overrightarrow{\theta}:f^{*}(\overrightarrow{\theta};y,x)=1\})=1.$$

**证** 令

$$\overrightarrow{\Theta}_{0}=\bigcap_{j,k\in\mathbf{Z}}\left\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:\begin{array}{c}f^{*}(T^{j}\overrightarrow{\theta};x,y)>0,\\f^{*}(T^{k}\overrightarrow{\theta};x,x)=1,\\q(\overrightarrow{\theta};x,x)=1\end{array}\right\},$$

则由 $\pi\circ T^{-1}=\pi$ 及命题 2.7 知 $\pi(\overrightarrow{\Theta}_{0})=1$.

对任一 $\overrightarrow{\theta}\in\overrightarrow{\Theta}_{0}$, 令

$$M(\overrightarrow{\theta})\overset{\text{def.}}{=\!=}\{f^{(s)}(\overrightarrow{\theta};x,y)=0,\forall_{1}\leqslant s<n,f^{(n)}(\overrightarrow{\theta};x,y)>0\},$$

显然,

$$M:\overrightarrow{\Theta}_{0}\mapsto\{1,2,\cdots\}$$

且关于 $\sigma$ 代数 $\overrightarrow{\Theta}_{0}\bigcap\overrightarrow{\mathscr{B}}$ 是可测的.

(注意: 本书恒用下列记法, 对 $\sigma$ 代数 $\mathscr{G}$ 中任一集合 $A\in\mathscr{G},A\bigcap\mathscr{G}$ 定义为 $\{B=A\bigcap C:C\in\mathscr{G}\}$.)

今 $x,y\in\mathscr{X}$ 是固定的, 而 $M(\overrightarrow{\theta})$ 有时简记为 $M$. 记

$$\alpha(\overrightarrow{\theta})\overset{\text{def.}}{=\!=}f^{(M(\overrightarrow{\theta}))}(\overrightarrow{\theta};x,y)=f^{(M)}(\overrightarrow{\theta};x,y),$$

则由 (1.17), 对任何 $\vec{\theta} \in \vec{\Theta}_0$ 有

$$\begin{aligned}
0 < \alpha(\vec{\theta}) &= f^{(M)}(\vec{\theta}; x, y) \\
&= P_{(x,\vec{\theta})}(X_s \neq y, \forall 1 \leqslant s < M, X_M = y) \\
&\leqslant P_{(x,\vec{\theta})}\left(X_s \neq y, \forall 1 \leqslant s \leqslant M, X_M = y, \bigcup_{n>M} \{X_n = x\}\right) \\
&\quad + P_{(x,\vec{\theta})}\left(X_M = y, \bigcap_{n>M} \{X_n \neq x\}\right) \\
&\overset{(1.17)}{=\!=} f^{(M)}(\vec{\theta}; x, y) P_{(y, T^M \vec{\theta})}\left(\bigcup_{n=1}^{\infty} \{X_n = x\}\right) \\
&\quad + P_{(x,\vec{\theta})}\left(X_M = y, \bigcap_{n>M} \{X_n \neq x\}\right) \\
&\leqslant f^{(M)}(\vec{\theta}; x, y) f^*(T^M \vec{\theta}; y, x) \\
&\quad + P_{(x,\vec{\theta})}\left(\bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{X_n \neq x\}\right) \\
&= \alpha(\vec{\theta}) f^*(T^M \vec{\theta}; y, x) + (1 - q(\vec{\theta}; x, x)) \\
&= \alpha(\vec{\theta}) f^*(T^M \vec{\theta}; y, x),
\end{aligned}$$

而 $\alpha(\vec{\theta}) > 0$, 由上式知

$$f^*(T^M \vec{\theta}; y, x) = 1 (\forall \vec{\theta} \in \vec{\Theta}_0). \tag{2.31}$$

但是对每个 $\vec{\theta} \in \vec{\Theta}_0$, 有

$$T^{M(\vec{\theta})} \vec{\Theta}_0 \supset \bigcap_{k \in \mathbf{Z}} T^k \vec{\Theta}_0,$$

而 $\pi = \pi \circ T^{-1}, \pi(\vec{\Theta}_0) = 1$, 故 $\pi(T^k \vec{\Theta}_0) = 1$, 从而

$$\pi(T^{M(\vec{\theta})} \vec{\Theta}_0) = 1 \quad (\forall \vec{\theta} \in \vec{\Theta}_0). \tag{2.32}$$

由 (2.31) 和 (2.32) 即得命题 2.12.

**命题 2.13** 若 $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, y) = \infty\}) = 1$, 则存在一个 $\vec{\mathscr{B}}$ 可测函数 $s = s(\vec{\theta})$, 使

$$p^{(s(\vec{\theta}))}(\vec{\theta}; x, y) > 0, \quad \pi - \text{a.s.}. \tag{2.33}$$

**证** 因为

$$\pi(\{\vec{\theta} \in \vec{\Theta} : \sum_{n=0}^{\infty} p^{(n)}(\vec{\theta}; x, y) = \infty\}) = 1,$$

$$p^{(n)}(\vec{\theta}; x, y) \text{ 是 } \vec{\theta} \text{ 的 } \mathscr{B} \text{ 可测函数,}$$

所以, 若令

$$s(\vec{\theta}) = \min\{n \geqslant 1 : p^{(u)}(\vec{\theta}; x, y) = 0, \forall 1 \leqslant u < n, p^{(n)}(\vec{\theta}; x, y) > 0\},$$

则 $s = s(\vec{\theta})$ 即为所求.

**附注 2.1** 若 $\pi = \pi \circ T^{-1}$, 则由命题 2.11 和 Borel – Contelli 引理知:

$$\begin{aligned} &\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, x) = 1\}) = 1 \\ \Leftrightarrow &\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; x, x) = 1\}) = 1 \\ \Rightarrow &\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, x) = \infty\}) = 1. \end{aligned}$$

但是, $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, x) = \infty\}) = 1$ 未必蕴涵了 $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, x) = 1\}) = 1$.

反例如下: 取 $\mathscr{X} = \{a, b\}, a \neq b, \Theta = \{0, 1\}, \Phi(A \times B) = \nu(A)\pi(B), \nu$ 是 $\mathscr{X}$ 上的离散布, $\pi = \gamma^{\mathbf{Z}}, \gamma(\{0\}) = \gamma(\{1\}) = \dfrac{1}{2}$, 则易证 $\pi \circ T^{-1} = \pi, \pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; a, a) = \infty\}) = 1$, 但 $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; a, a) < 1\}) = 1$.

**定理 2.1** 设 $\pi = \pi \circ T^{-1}$, 且 $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, y) > 0\}) = 1, \pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, x) = 1\}) = 1$, 则

(1) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, x) = 1\}) = 1$;

(2) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; y, x) = 1\}) = 1$;

(3) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, y) = 1\}) = 1$;

(4) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; y, y) = 1\}) = 1$;

(5) 把结论 (1) ~ (4) 中的 $f^*_{\min}$ 换成 $q_{\min}$, 上面四个相应的结论仍然成立.

**证** 因为

$$\begin{aligned} &\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; u, v) = 1\} \\ = &\bigcap_{k \in \mathbf{Z}} \{\vec{\theta} \in \vec{\Theta} : f^*(T^k \vec{\theta}; u, v) = 1\}, \end{aligned}$$

所以, (1) 可由 $\pi = \pi \circ T^{-1}$ 直接得到, 而 (2) 可以由 $\pi = \pi \circ T^{-1}$ 和命题 2.12 得到.

(3) 由命题 2.5 知: $f^*(\vec{\theta};x,y) \geqslant q(\vec{\theta};x,y) \geqslant q(\vec{\theta};x,x)$, $\pi-$a.s..
由命题 2.11, 有

$$\begin{aligned}
&\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta};x,y)=1\}) \\
\geqslant &\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta};x,x)=1\}) \\
= &\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta};x,x)=1\}) \\
= &1.
\end{aligned}$$

仿 (2), 用 $\pi=\pi\circ T^{-1}$ 可证 (3).

(4) $f^*(\vec{\theta};y,y)$

$$\begin{aligned}
&\geqslant P_{(y,\vec{\theta})}\left(\bigcup_{n=1}^{\infty}\{X_n=y\},\bigcup_{n=1}^{\infty}\{X_n=x\}\right) \\
&= f^*(\vec{\theta};y,x)-\sum_{n=1}^{\infty}P_{(y,\vec{\theta})}\left(X_k\neq x, 1\leqslant k<n, X_n=x, \bigcap_{s=n+1}^{\infty}\{X_s\neq y\}\right) \\
&= f^*(\vec{\theta};y,x)-\sum_{n=1}^{\infty}f^{(n)}(\vec{\theta};y,x)(1-f^*(T^n\vec{\theta};x,y)) \\
&= f^*(\vec{\theta};y,x), \quad \pi-\text{a.s.}.
\end{aligned}$$

所以由 $\pi=\pi\circ T^{-1}$ 及上式可得 (4).

(5) 由 (1)~(4) 及 (2.29)* 立即可证得 (5) 成立.

**推论 2.1**　*在定理 2.1 的条件下, 恒有:*

$$\begin{aligned}
&\pi(\{\vec{\theta} \in \vec{\Theta} : g_{\min}(\vec{\theta};x,y)=\infty\}) \\
= &\pi(\{\vec{\theta} \in \vec{\Theta} : g_{\min}(\vec{\theta};x,x)=\infty\}) \\
= &\pi(\{\vec{\theta} \in \vec{\Theta} : g_{\min}(\vec{\theta};y,x)=\infty\}) \\
= &\pi(\{\vec{\theta} \in \vec{\Theta} : g_{\min}(\vec{\theta};y,y)=\infty\}) \\
= &1.
\end{aligned}$$

**证**　由 Borel – Contelli 引理知:

$$“q(\vec{\theta};x,y)=P_{(x,\vec{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=y\}\right)>0”$$

蕴涵了

$$“G(\vec{\theta};x,y)=\sum_{k=0}^{\infty}P_{(x,\vec{\theta})}(X_k=y)=\infty”.$$

所以, 推论 2.1 可以由 $\pi=\pi\circ T^{-1}$ 和定理 2.1 (5) 推出.

**推论 2.2** 若 $\pi=\pi\circ T^{-1}, \pi(\{\vec{\theta}\in\vec{\Theta}: f^*(\vec{\theta};x,x)=1\})=1$, 且存在 $\varepsilon>0$ 使

$$\pi(\{\vec{\theta}: f^*(\vec{\theta};x,y)\geqslant\varepsilon\})=1,$$

则定理 2.1 和推论 2.1 的结论全部成立.

**证** 由 "$\pi=\pi\circ T^{-1}$ 和 $\pi(\{\vec{\theta}: f^*(\vec{\theta};x,y)\geqslant\varepsilon\})=1$" 可以推出

$$\pi(\{\vec{\theta}\in\vec{\Theta}: f^*_{\min}(\vec{\theta};x,y)>0\})=1.$$

因此, 推论 2.2 可以由定理 2.1 和推论 2.1 而得.

## §3 状态的分类

在第六章的经典的 Markov 链中, 我们已经看到状态的分类的重要性, 它对 Markov 链的理论研究, 特别是对 Markov 链的极限理论的研究, 有重要意义. 自然地, 对随机环境中的 Markov 链, 我们也要研究状态的分类. 但它比经典 Markov 链的状态分类要复杂得多

本节沿袭前二节的符号. 如不特别声明, 各种符号的意义均与前二节相同.

还需强调指出: $(\vec{X},\vec{\xi}), \{(X_n, T\vec{\xi}); n\geqslant 0\}$ 分别是定理 1.1 中所构造的 MCRE 和 SPMC, $p(\cdot)$ 是 MCRE 的随机转移矩阵, $p^{(n)}(\vec{\theta};x,y)$ 如 (1.2) 所定义, $P^x_{\vec{\theta}}(\cdot)$ 是由 (1.7) 所定义的 $\vec{\mathscr{A}}$ 上的概率测度, $\widehat{P}^{(n)}((x,\vec{\theta}),F)$ 是由 (1.8) 所定义的转移函数, $P_{(x,\vec{\theta})}(\cdot)$ 是由 (1.12) 所定义的 $\mathscr{A}\times\vec{\mathscr{B}}$ 上的概率测度. $p^{(n)}(\vec{\theta};x,y)$ 和 $\widehat{P}^{(n)}((x,\vec{\theta}),F)$ 的概率意义分别是:

$$p^{(n)}(\vec{\theta};x,y)=P_{(x,\vec{\theta})}(X_n=y);$$
$$\widehat{P}^{(n)}((x,\vec{\theta}),F)=P_{(x,\vec{\theta})}((X_n,T^n\vec{\xi})\in F).$$

下面回顾一些常用的符号的定义.

$F^{(n)}((x,\vec{\theta}),G)$, $f^{(n)}(\vec{\theta};x,y)$, $\widehat{F}^*((x,\vec{\theta}),G)$, $f^*(\vec{\theta};x,y)$, $Q((x,\vec{\theta}),G)$, $q(\vec{\theta},x,y)$ 的定义分别见 (2.4)、(2.5)、(2.6)、(2.7)、(2.8) 和 (2.9).

本节恒设 $\pi=\pi\circ T^{-1}$.

**定义 3.1** 固定任何 $x,y\in\mathscr{X}$, 称 $x$ **可达** $y$, 记作 $x\to y$, 如果 $\pi(\{\vec{\theta}\in\vec{\Theta}: f^*(\vec{\theta};x,y)>0\})=1$. 称 $x$ **齐性可达** $y$, 记作 $x\xrightarrow{\min}y$, 如果 $\pi(\{\vec{\theta}\in\vec{\Theta}: f^*_{\min}(\vec{\theta};x,y)>0\})=1$. 称 $x$ **正性可达** $y$, 记作 $x\xrightarrow{\mathrm{p}}y$, 如果存在 $\varepsilon>0$, 使 $\pi(\{\vec{\theta}\in\vec{\Theta}: f^*(\vec{\theta};x,y)\geqslant\varepsilon\})=1$.

**命题 3.1** 对任何 $x,y\in\mathscr{X}$, 恒有

$$x\xrightarrow{\mathrm{p}}y\Rightarrow x\xrightarrow{\min}y\Rightarrow x\to y.$$

证明甚易, 读者可作为习题验证之.

**命题 3.2**　若 $x \to y, y \to z$, 则 $x \to z$.

**证**　由 $x \to y, y \to z$ 得

$$\begin{aligned}&\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, y) > 0\}) \\ =&\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; y, z) > 0\}) \\ =&1.\end{aligned}$$

若令

$$\vec{\Theta}_0 = \bigcap_{j,k\in\mathbf{Z}} \{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, y) f^*(\vec{\theta}; y, z) > 0\},$$

则由 $\pi = \pi \circ T^{-1}$ 可知

$$\pi(\vec{\Theta}_0) = 1.$$

任取 $\vec{\theta} \in \vec{\Theta}_0$, 存在 $u = u(\vec{\theta}), v = v(\vec{\theta})$ 使

$$f^{(u)}(\vec{\theta}; x, y) > 0, \quad f^{(v)}(T^u\vec{\theta}; y, z) > 0.$$

更有

$$p^{(u)}(\vec{\theta}; x, y) > 0, \quad p^{(v)}(T^u\vec{\theta}; y, z) > 0.$$

由 (1.2)、(2.3) 和 (1.17) 得

$$\begin{aligned}&p^{(u+v)}(\vec{\theta}; x, z) \geqslant P_{(x,\vec{\theta})}(X_u = y, X_{u+v} = z) \\ =&P_{(x,\vec{\theta})}(X_u = y) P_{(y,T^u\vec{\theta})}(X_v = z) \\ =&p^{(u)}(\vec{\theta}; xy) p^{(v)}(T^u\vec{\theta}; y, z) > 0.\end{aligned}$$

因此, 由命题 2.3 知: 存在 $w \leqslant u + v$, 使

$$f^{(w)}(\vec{\theta}; x, z) > 0 \quad (\forall \vec{\theta} \in \vec{\Theta}_0).$$

更有

$$f^*(\vec{\theta}; x, z) > 0 \quad (\forall \vec{\theta} \in \vec{\Theta}_0).$$

由 $\pi(\vec{\Theta}_0) = 1$ 得 $x \to z$.

**命题 3.3**　若 $x \xrightarrow{\min} y, y \xrightarrow{\min} z$, 则 $x \xrightarrow{\min} z$.

**证**　设 $x \xrightarrow{\min} y, y \xrightarrow{\min} z$. 令

$$\vec{\Theta}_0 = \{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, y) f^*_{\min}(\vec{\theta}; y, z) > 0\}$$

则 $\pi(\overrightarrow{\Theta}_0)=1$, 且对任何 $\overrightarrow{\theta}\in\overrightarrow{\Theta}_0$, 存在 $\varepsilon=\varepsilon(\overrightarrow{\theta})>0$, 使

$$f^*(T^j\overrightarrow{\theta};x,y)\geqslant\varepsilon\quad(\forall j\in\mathbf{Z}),\quad f^*(T^k\overrightarrow{\theta};y,z)\geqslant\varepsilon\quad(\forall k\in\mathbf{Z}).$$

但是对任何 $r\in\mathbf{Z}$, 由 (1.17) 有

$$\begin{aligned}
f^*(T^r\overrightarrow{\theta};x,z)&=P_{(x,\Gamma^r\overrightarrow{\theta})}\left(\bigcup_{n=1}^{\infty}\{X_n=z\}\right)\\
&\geqslant P_{(x,T^r\overrightarrow{\theta})}\left(\bigcup_{s=1}^{\infty}\{X_s=y\},\bigcup_{n=1}^{\infty}\{X_n=z\}\right)\\
&=\lim_{l\to\infty}\lim_{k\to\infty}P_{(x,T^r\overrightarrow{\theta})}\left(\bigcup_{s=1}^{l}\{X_s=y\},\bigcup_{n=1}^{k}\{X_n=z\}\right)\\
&\geqslant\lim_{l\to\infty}\lim_{k\to\infty}\sum_{s=1}^{l}P_{(x,T^r\overrightarrow{\theta})}\left(X_u\neq y,1\leqslant u<s,X_s=y,\bigcup_{n=s+1}^{k}\{X_n=z\}\right)\\
&\overset{(1.17)}{=\!=}\lim_{l\to\infty}\lim_{k\to\infty}\sum_{s=1}^{l}f^{(s)}(T^r\overrightarrow{\theta};x,y)\\
&\quad\cdot P_{(y,T^{s+r}\overrightarrow{\theta})}\left(\bigcup_{n=1}^{k-s}\{X_n=z\}\right)\\
&=\lim_{l\to\infty}\lim_{k\to\infty}\sum_{s=1}^{l}f^{(s)}(T^r\overrightarrow{\theta};x,y)\sum_{n=1}^{k-s}f^{(n)}(T^{s+r}\overrightarrow{\theta};y,z)\\
&=\lim_{l\to\infty}\sum_{s=1}^{l}f^{(s)}(T^r\overrightarrow{\theta};x,y)f^*(T^{s+r}\overrightarrow{\theta};y,z)\\
&\geqslant\varepsilon^2\quad(\forall\overrightarrow{\theta}\in\overrightarrow{\Theta}_0).
\end{aligned}$$

所以

$$f^*_{\min}(\overrightarrow{\theta};x,z)>0\quad(\forall\overrightarrow{\theta}\in\overrightarrow{\Theta}_0),$$

故 $x\xrightarrow{\min}z$.

**命题 3.4** 若 $x\xrightarrow{\mathrm{p}}y,y\xrightarrow{\mathrm{p}}z$, 则 $x\xrightarrow{\mathrm{p}}z$.

**证** 设 $x\xrightarrow{\mathrm{p}}y,y\xrightarrow{\mathrm{p}}z$, 则存在 $\varepsilon>0$ 不依赖 $\overrightarrow{\theta}$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:f^*(\overrightarrow{\theta};x,y)f^*(\overrightarrow{\theta};y,z)\geqslant 0\})=1.$$

令

$$\overrightarrow{\Theta}_1=\bigcap_{j,k\in\mathbf{Z}}\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:f^*(T^j\overrightarrow{\theta};x,y)\geqslant\varepsilon,f^*(T^k\overrightarrow{\theta};y,z)\geqslant\varepsilon\}$$

则由 $\pi = \pi \circ T^{-1}$ 得 $\pi(\vec{\Theta}_1) = 1$. 任取 $\vec{\theta} \in \vec{\Theta}_1$, 存在 $u = u(\vec{\theta})$ 和 $v = v(\vec{\theta})$, 使

$$P_{(x,\vec{\theta})}\left(\bigcup_{s=1}^{u}\{X_s = y\}\right) \geqslant \frac{\varepsilon}{2},$$
$$P_{(y,T^s\vec{\theta})}\left(\bigcup_{l=1}^{v}\{X_l = z\}\right) \geqslant \frac{\varepsilon}{2} \quad (\forall 1 \leqslant s \leqslant u).$$

因此, 由 (1.17) 得

$$\begin{aligned}
&P_{(x,\vec{\theta})}\left(\bigcup_{l=1}^{u+v}\{X_l = z\}\right) \geqslant P_{(x,\vec{\theta})}\left(\bigcup_{s=1}^{u}\{X_s = y\}, \bigcup_{l=1}^{u+v}\{X_l = z\}\right) \\
= \;& \sum_{s=1}^{u} P_{(x,\vec{\theta})}\left(X_k \neq y, 1 \leqslant k < s, X_s = y, \bigcup_{l=1}^{u+v}\{X_l = z\}\right) \\
\geqslant \;& \sum_{s=1}^{u} P_{(x,\vec{\theta})}\left(X_k \neq y, 1 \leqslant k < s, X_s = y, \bigcup_{l=s+1}^{u+v}\{X_l = z\}\right) \\
\xlongequal{(1.17)} \;& \sum_{s=1}^{u} f^{(s)}(\vec{\theta}; x, y) P_{(y,T^s,\vec{\theta})}\left(\bigcup_{l=1}^{u+v-s}\{X_l = z\}\right) \\
\geqslant \;& \frac{\varepsilon^2}{4} \quad (\forall \vec{\theta} \in \vec{\Theta}_1).
\end{aligned}$$

所以

$$\pi\left(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, z) \geqslant \frac{\varepsilon^2}{4}\right) \geqslant \pi(\vec{\Theta}_1) = 1,$$

从而 $x \xrightarrow{\mathrm{p}} z$ 命题证毕.

**定义 3.2** 称状态 $x$ 和 $y$ 是**互通的**, 记作 $x \leftrightarrow y$, 如果 $x \to y$ 而且 $y \to x$. 仿之可定义 $x \xleftrightarrow{\min} y$ 和 $x \xleftrightarrow{\mathrm{p}} y$.

**定义 3.3** 称状态 $x$ 是**非本质的**, 如果对任何 $y \in \mathscr{X}$, 有 $\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; y, x) = 0\}) = 1$; 反之, 称 $x$ 是**本质的**.

称状态 $x$ 是**常返的**, 若 $\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; x, x) = 1\}) = 1$; 反之, 称 $x$ 是暂留的.

称常返状态 $x$ 是**正常返的**, 如果

$$\begin{aligned}
&\pi\left(\{\vec{\theta} \in \vec{\Theta} : \liminf_{\lambda \uparrow 1}(1-\lambda) m_p(T^k\vec{\theta}; x, x, \lambda) > 0\}\right) \\
&= 1 \quad (\forall k \geqslant 0),
\end{aligned} \tag{3.1}$$

其中 $m_p(\vec{\theta}; x, y, \lambda)$ 之定义见 (2.15). 若 (3.1) 不成立, 则称 $x$ 是零常返的.

**命题 3.5** 若对任何 $n \geqslant 0$, $k \in \mathbf{Z}$, $x,y \in \mathscr{X}$, 有 $p^{(n)}(T^k\vec{\theta};x,y) = p^{(n)}(\vec{\theta};x,y)$, 则

$$\liminf_{\lambda\uparrow 1}(1-\lambda)m_p(\vec{\theta};x,x,\lambda) = \left(\sum_{n=1}^{\infty} nf^{(n)}(\vec{\theta};x,x)\right)^{-1} \tag{3.2}$$

证明甚易, 读者可作为习题验证之.

**命题 3.6** 若 $x$ 是正常返的, 且 $x \xrightarrow{\min} y$, 则 $y$ 也是正常返的.

**证** 根据命题 3.6 的假设, 定理 2.1 的条件全部满足, 仿定理 2.1 之证, 存在两个正的 $\vec{\mathscr{B}}$ 可测函数 $s = s(\vec{\theta})$ 和 $t = t(T^{s+n}(\vec{\theta}))$, 使

$$p^{(s)}(\vec{\theta};y,x) > 0, \quad p^{(t)}(T^{s+n}(\vec{\theta}),x,y) > 0 \tag{3.3}$$

$(\forall \vec{\theta} \in \vec{\Theta}_0, \pi(\vec{\Theta}_0) = 1)$. 由定理 2.1 知 $y$ 是常返的.

由于 $x$ 是正常返的, 所以又存在 $\vec{\Theta}_1$, 使 $\pi(\vec{\Theta}_1) = 1$, 且对任何 $\vec{\theta} \in \vec{\Theta}_1$, 任何 $k \geqslant 0$, 有

$$\liminf_{\lambda\uparrow 1}(1-\lambda)m_p(T^k\vec{\theta};x,x,\lambda) > 0. \tag{3.4}$$

注意: 用 (1.17) 可证;

$$\begin{aligned}
m_p(T^k\vec{\theta};y,y,\lambda) &= \sum_{n=0}^{\infty} p^{(n)}(T^k\vec{\theta};y,y)\lambda^n \\
\geqslant & \sum_{n=s+t}^{\infty} p^{(n)}(T^k\vec{\theta};y,y)\lambda^n \\
= & \sum_{n=s+t}^{\infty} P_{(y,T^k\vec{\theta})}(X_n = y)\lambda^n \\
\geqslant & \sum_{n=s+t}^{\infty} P_{(y,T^k\vec{\theta})}(X_s = x, X_{n-t} = x, X_n = y)\lambda^n \\
\overset{(1.17)}{=\!=} & \sum_{n=s+t}^{\infty} P_{(y,T^k\vec{\theta})}(X_s = x)P_{(x,T^{k+s}\vec{\theta})}(X_{n-t-s} = x, X_{n-s} = y)\lambda^n \\
\overset{(1.17)}{=\!=} & \sum_{n=s+t}^{\infty} P_{(y,T^k\vec{\theta})}(X_s = x)P_{(x,T^{k+s}\vec{\theta})}(X_{n-t-s} = x) \\
& \cdot P_{(x,T^{k+s+t}\vec{\theta})}(X_t = y)\lambda^n \\
= & \sum_{n=s+t}^{\infty} \Big[p^{(s)}(T^k\vec{\theta};y,x)\lambda^s \cdot p^{(n-t-s)}(T^{k+s}\vec{\theta};x,x)\lambda^{n-t-s} \\
& \cdot p^{(t)}(T^{k+s+t}\vec{\theta};x,y)\lambda^t\Big] \\
= & \ p^{(s)}(T^k\vec{\theta};y,x)\lambda^s(m_p(T^{k+s}\vec{\theta};x,x,\lambda)) \cdot p^{(t)}(T^{k+s+t}\vec{\theta};x,y)\lambda^t. \qquad (3.5)
\end{aligned}$$

由 (3.3)、(3.4)、(3.5) 可知: 若取 $\vec{\Theta}_2 = \vec{\Theta}_1 \bigcap \vec{\Theta}_0$, 则 $\pi(\vec{\Theta}_2) = 1$, 且对任何 $\vec{\theta} \in \vec{\Theta}_2, k \geqslant 0$, 都有

$$\liminf_{\lambda \uparrow 1} (1-\lambda) m_p(T^k \vec{\theta}; y, y, \lambda) > 0,$$

此即 $y$ 是正常返状态. 命题证毕.

**定义 3.4** 称状态集合 $J \subset \mathscr{X}$ 是**封闭的**, 如果 $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, y) = 0\}) = 1$ 对一切 $x \in J$ 和 $y \overline{\in} J$ 成立. 称 $J$ 是**弱封闭的**, 如果 $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, y) = 0\}) > 0$ 对一切 $x \in J$ 和 $y \overline{\in} J$ 成立. 称 $J$ 是 $M$ **封闭的**, 如果 $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, y) = 0\}) > 0$ 对一切 $x \in J, y \overline{\in} J$ 成立.

**命题 3.7** 恒有下列诸结论:

(1) $J$ 是封闭的 $\Rightarrow$ $J$ 是弱封闭的 $\Rightarrow$ $J$ 是 M 封闭的.

(2) 下列陈述等价:

(i) $J$ 是封闭的;

(ii) 存在集合 $\vec{\Theta}_1$ 使 $\pi(\vec{\Theta}_1) = 1$, 而且 $p^{(n)}(T^k \vec{\theta}; x, y) = 0$ $(\forall n \geqslant 0, x \in J, y \overline{\in} J, k \in \mathbf{Z}, \vec{\theta} \in \vec{\Theta}_1)$;

(iii) 存在集合 $\vec{\Theta}_1$, 使 $\pi(\vec{\Theta}_1) = 1$, 而且 $f^*(T^k \vec{\theta}; x, y) = 0$ $(\forall x \in J, y \overline{\in} J, k \in \mathbf{Z}, \vec{\theta} \in \vec{\Theta}_1)$.

(3) 下列陈述等价:

(i) $J$ 是弱封闭的;

(ii) 对任何 $x \in J, y \overline{\in} J$, “$x \to y$” 不能成立 (即 $x$ 不可达 $y$);

(4) $J$ 是 $M$ 封闭的 $\Leftrightarrow$ 对任何 $x \in J, y \overline{\in} J$, “$x \xrightarrow{\min} y$” 不能成立.

证明甚易, 诸者可作为习题验证之.

**定理 3.1** 恒有下述诸结论:

(1) 若 $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, x) < \infty\}) = 1$, 则 $x$ 是非本质的;

(2) 下列陈述等价:

(i) 状态 $x$ 是常返的;

(ii) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*(\vec{\theta}; x, x) = 1\}) = 1$;

(iii) $\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; x, x) = 1\}) = 1$;

(iv) $\pi(\{\vec{\theta} \in \vec{\Theta} : f^*_{\min}(\vec{\theta}; x, x) = 1\}) = 1$;

(v) $\pi(\{\vec{\theta} \in \vec{\Theta} : q_{\min}(\vec{\theta}; x, x) = 1\}) = 1$.

(3) 下列陈述等价:

(i) $x$ 是非本质的;

(ii) $\pi\left( \bigcap_{y \in \mathscr{X}} \{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; y, x) = 0\} \right) = 1$;

(iii) $\pi(\{\vec{\theta}\in\vec{\Theta}:q(\vec{\theta};x,x)=0\})=1$.

(4) 若 $x$ 是非本质, 则 $x$ 是暂留的.

**证** (1) 由于

$$g(\vec{\theta};y,x)=\sum_{n=0}^{\infty}p^{(n)}(\vec{\theta};y,x)$$
$$=\sum_{n=0}^{\infty}P_{(y,\vec{\theta})}(X_n=x),$$

所以由假设知

$$\pi\left(\left\{\vec{\theta}\in\vec{\Theta}:\sum_{n=0}^{\infty}P_{(x,\vec{\theta})}(X_n=x)<\infty\right\}\right)=1. \tag{3.6}$$

由 Borel – Contelli 引理和 (3.6) 知

$$\pi(\{\vec{\theta}\in\vec{\Theta}:q(\vec{\theta};x,x)=0\})$$
$$=\pi\left(\left\{\vec{\theta}\in\vec{\Theta}:P_{(x,\vec{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=x\}\right)=0\right\}\right)=1. \tag{3.7}$$

但是

$$\begin{aligned}
q(\vec{\theta};y,x)&=P_{(y,\vec{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=x\}\right)\\
&=\sum_{l=1}^{\infty}P_{(y,\vec{\theta})}\left(X_u\neq x,1\leqslant u<l,X_l=x,\bigcap_{n=l+1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=x\}\right)\\
&\overset{(1.17)}{=\!=}\sum_{l=1}^{\infty}f^{(l)}(\vec{\theta};y,x)P_{(x,T^l\vec{\theta})}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}\{X_k=x\}\right)\\
&=\sum_{l=1}^{\infty}f^{(l)}(\vec{\theta};y,x)q(T^l\vec{\theta};x,x)\quad(\forall g\in\mathscr{X}).
\end{aligned}\tag{3.8}$$

由 $\pi=\pi\circ T^{-1}$ 及 (3.7) 得知: 存在 $\vec{\Theta}_0$, 使 $\pi(\vec{\Theta}_0)=1$ 且

$$q(T^l\vec{\theta};x,x)=0\ \ (\forall l\geqslant 1,\vec{\theta}\in\vec{\Theta}_0). \tag{3.9}$$

以 (3.9) 代入 (3.8) 得

$$q(\vec{\theta};y,x)=0\ \ (\forall\vec{\theta}\in\vec{\Theta}_0,y\in\mathscr{X}).$$

而 $\pi(\vec{\Theta}_0)=1$, 所以 $x$ 是非本质的.

(2) 可由命题 2.11 而得.

(3) 在 (1) 的证明中已经证明了: 由

$$\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; x, x) = 0\}) = 1$$

可推出

$$\pi(\{\vec{\theta} \in \vec{\Theta} : q(\vec{\theta}; y, x) = 0\}) = 1 \quad (\forall y \in \mathscr{X}),$$

所以 (3) 成立.

(4) 由 (2) 和 (3) 可推出 (4). 定理证毕.

## §4 状态的周期及状态空间的分解

本节仍沿袭前三节的符号, 而且恒设:

(1) $\pi = \pi \circ T^{-1}$;

(2) $0-1$ 律条件成立: 对任何正整数 $n$, 任何 $x, y \in \mathscr{X}$, 总有 $\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(n)}(\vec{\theta}; x, y) > 0\}) = 0$ 或 1.

**定义 4.1** 设状态 $x$ 满足: $\pi(\{\vec{\theta} \in \vec{\Theta} : g(\vec{\theta}; x, x) = \infty\}) = 1$. 称

$$\text{G.C.D.}\{n \geqslant 1 : \pi(\{\vec{\theta} \in \vec{\Theta} : p^{(n)}(\vec{\theta}; x, x) > 0\}) = 1\}$$

为 $x$ 之**周期**, 记之为 $d_x$, 其中 G.C.D. 表示最大公因子.

注意: 常返状态 $x$ 的周期总是有定义的.

**定理 4.1** 设 $x$ 是常返状态, 且 $x \xrightarrow{\min} y$, 则 $d_x = d_y$.

(注意: 在定理的条件下, $y$ 也是常返状态, 而 $y \xrightarrow{\min} x$. 所以 $x$ 和 $y$ 的地位是完全对称的.)

**证** 为证定理, 只需证明

$$\text{“}\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(n)}(\vec{\theta}; y, y) > 0\}) = 1 \Rightarrow d_x | n\text{”}.$$

(此处及今后 $k|n$ 表示 $k$ 能整除 $n$.)

事实上若上述蕴涵关系成立, 则 $d_x | d_y$. 由于 $d_x$ 和 $d_y$. 地位的对称性, 亦有 $d_y | d_x$, 从而 $d_x = d_y$, 下面我们来证明上述蕴涵关系.

设

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(n)}(\vec{\theta}; y, y) > 0\}) = 1. \tag{4.1}$$

由命题 2.13 和推论 2.1 和 $0-1$ 条件可知: 存在正整数 $s$ 和 $t$ (不依赖 $\vec{\theta}$) 使

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(s)}(\vec{\theta}; x, y) > 0\}) = 1, \tag{4.2}$$

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(t)}(T^{s+n}\vec{\theta}; y, x) > 0\}) = 1. \tag{4.3}$$

但是, 由 (1.17) 知

$$p^{(s+n+t)}(\vec{\theta};x,x) \geqslant \left[p^{(s)}(\vec{\theta};x,y)\bullet p^{(n)}(T^s\vec{\theta};y,y)p^{(t)}(T^{s+n}\vec{\theta};y,x)\right]. \tag{4.4}$$

由 (4.1)~(4.4) 可推出:

$$\pi(\{\vec{\theta}\in\vec{\Theta}:p^{(s+n+t)}(\vec{\theta};x,x)>0\})=1. \tag{4.5}$$

所以 $d_x|(s+n+t)$.

但是, 由 $\pi=\pi\circ T^{-1}$, (4.1) 及

$$p^{(2n)}(\vec{\theta};y,y) \geqslant p^{(n)}(\vec{\theta};y,y)p^{(n)}(T^n\vec{\theta};y,y). \tag{4.6}$$

可推出

$$\pi(\{\vec{\theta};\in\vec{\Theta}:p^{(2n)}(\vec{\theta};y,y)>0\})=1. \tag{4.7}$$

而

$$p^{(s+2n+t)}(\vec{\theta};x,x) \geqslant \left[p^{(s)}(\vec{\theta};x,y)\bullet p^{(2n)}(T^s\vec{\theta};y,y)p^{(t)}(T^{s+2n}\vec{\theta};y,x)\right], \tag{4.8}$$

仿前可证

$$\pi(\{\vec{\theta}\in\vec{\Theta}:p^{(s+2n+t)}(\vec{\theta};x,x)>0\})=1. \tag{4.9}$$

所以 $d_x|(s+2n+t)$, 从而 $d_x|n$. 定理证毕.

**定理 4.2** 令 $\mathscr{N}\subset\mathscr{X}$ 为全体暂留状态, $\mathscr{R}$ 为全体常返状态, $\mathscr{R}_+$ 为全体正常返状态, $\mathscr{R}_0$ 为全体零常返状态, 则

$$\mathscr{X}=\mathscr{N}\bigcup\mathscr{R}=\mathscr{N}\bigcup(\mathscr{R}_0\bigcup\mathscr{R}_+), \tag{4.10}$$

而且

(1) $x\in\mathscr{R},y\in\mathscr{N}\Rightarrow$ "$x\overset{\min}{\longrightarrow}y$ 不可能发生;"

(2) $x\in\mathscr{R}_0,y\in\mathscr{R}_+\Rightarrow$ "$x\overset{\min}{\longrightarrow}y$ 和 $y\overset{\min}{\longrightarrow}x$ 都不可能发生";

**证** (1) 由定理 2.1 知: 当 $x\in\mathscr{R}$, 且 $x\overset{\min}{\longrightarrow}y$ 时, $y$ 必为常返状态, 所以 (1) 成立.

(2) 当 $y\in\mathscr{R}_+$ 且 $y\overset{\min}{\longrightarrow}x$ 时, 由命题 3.6 知 $x\in\mathscr{R}_+$. 因此 $x\in\mathscr{R}_0$ 和 $y\in\mathscr{R}_+$ 蕴涵了 $y\overset{\min}{\longrightarrow}x$ 不可能发生.

因为由定理 2.1 知: $x\in\mathscr{R}_0$ 和 $x\overset{\min}{\longrightarrow}y$ 蕴涵了 $y\overset{\min}{\longrightarrow}x$, 而今 $y\in\mathscr{R}_+$, 所以 $x\in\mathscr{R}_+$. 这与 $x\in\mathscr{R}_0$ 的假设矛盾, 所以 $x\overset{\min}{\longrightarrow}y$ 不可能发生.

由定理 2.1 和命题 3.3 及定理 4.2 易见: 关系 $\overset{\min}{\longleftrightarrow}$ 在 $\mathscr{R}_0$ 中 (在 $\mathscr{R}_+$ 中也一样) 是一个等价关系, 所以 $\mathscr{R}_0$($\mathscr{R}_+$ 也类似) 可以分解成一些等价类:

$$\mathscr{R}_0=\bigcup_{i\in\Gamma_0}\mathscr{R}_0(i),\quad \mathscr{R}_+=\bigcup_{j\in\Gamma_1}\mathscr{R}_+(j), \tag{4.11}$$

$x, y$ 同属于某个等价类 $\mathscr{R}_0(i)$ (或某个 $\mathscr{R}_+(j)$) 的充分必要条件是:

$$x \overset{\min}{\longleftrightarrow} y,$$

此处 $\Gamma_0$ 和 $\Gamma_1$ 是可数无穷集或有限集.

由定理 4.1 知: 周期 $d_x$ 不依赖于 $x$ 而只依赖于 $x$ 所在的某个等价类 $\mathscr{R}_0(i)$ (或 $\mathscr{R}_+(j)$). 所以, 当 $x \in \mathscr{R}_0(i)$ (或 $\mathscr{R}_+(j)$), 我们可以把 $d_x$ 写成 $d_0(i)$ (或 $d_+(j)$).

**定义 4.2** 对于任意的 $x, y \in \mathscr{R}_0(i)$, 令

$$H_{x,y} = \text{G.C.D.}\{n \geqslant 1 : \pi(\{\vec{\theta} \in \vec{\Theta} : p^{(n)}(\vec{\theta}; x, y) > 0\}) = 1\}, \tag{4.12}$$

若

$$d_0(i) | H_{x,y},$$

则称 $x$ 和 $y$ 具有关系 $\approx$, 记作 $x \approx y$.

**命题 4.1** 关系 $\approx$ 是 $\mathscr{R}_0(i)$ 中的一个等价关系.

**证** (1) 由于对任何 $x \in \mathscr{R}_0(i), d_0(i)|H_{x,x}$, 所以 $x \approx x$, 亦即关系 $\approx$ 有自返性.

(2) 对于任何 $x, y \in \mathscr{R}_0(i)$, 如果 $x \approx y$, 则 $\pi(\{\vec{\theta} : p^{(k)}(\vec{\theta}; x, y) > 0\}) = 1$ 蕴涵了 $d_0(i)|k$.

假定 $\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(l)}(\vec{\theta}; y, x) > 0\}) = 1$. 由 (1.17) 有:

$$p^{(k+l)}(\vec{\theta}; x, x) \geqslant p^{(k)}(\vec{\theta}; x, y) p^{(l)}(T^k \vec{\theta}; y, x),$$

所以, 用 $\pi = \pi \circ T^{-1}$ 可证

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(k+l)}(\vec{\theta}; x, x) > 0\}) = 1,$$

从而 $d_0(i)|(k+l)$. 所以 $d_0(i)|l$, 从而 $d_0(i)|H_{y,x}$. 因此, $y \approx x$. 这就证明了关系 $\approx$ 的对称性.

(3) 对于任何 $x, y, u \in \mathscr{R}_0(i)$, 若 $x \approx y, y \approx u$, 要证 $x \approx u$, 为此只需证对任何满足

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(r)}(\vec{\theta}; x, u) > 0\}) = 1 \tag{4.13}$$

的 $r$, 都有 $d_0(i)|r$. 由于 (2) 中已证关系 $\approx$ 有对称性, 故 $u \approx y, y \approx x$, $x, y, u \in \mathscr{R}_0(i)$ . 仿 (4.2) 和 (4.3) 式并注意 $\pi = \pi \circ T^{-1}$ 得知: 存在正整数 $s$ 和 $t$ 使

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(s)}(T^r \vec{\theta}; u, y) > 0\}) = 1, \tag{4.14}$$

$$\pi(\{\vec{\theta} \in \vec{\Theta} : p^{(t)}(T^{r+s} \vec{\theta}; y, x) > 0\}) = 1. \tag{4.15}$$

但是用 (1.17) 可证

$$p^{(r+s+t)}(\overrightarrow{\theta};x,x)\geqslant\left[p^{(r)}(\overrightarrow{\theta};x,u)\bullet p^{(s)}(T^{r}\overrightarrow{\theta};u,y)\bullet p^{(t)}(T^{r+s}\overrightarrow{\theta};y,x)\right]. \tag{4.16}$$

由 (4.13)~(4.16) 知

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{r+s+t}(\overrightarrow{\theta};x,x)>0\})=1. \tag{4.17}$$

由 $u\approx y,y\approx x,x\approx x$ 和 (4.14)、(4.15)、(4.17) 和 $\pi=\pi\circ T^{-1}$ 知:

$$d_0(i)|s,\quad d_0(i)|t,\quad d_0(i)|(r+s+t),$$

所以

$$d_0(i)|r. \tag{4.18}$$

这就证明了 $\approx$ 有传递性. 由 (1)、(2)、(3) 知关系 $\approx$ 是一个等价关系. 命题 4.1 证毕.

**命题 4.2** 对任何 $x,y\in\mathscr{R}_0(i)$ (当 $x,y\in\mathscr{R}_+(j)$ 时亦有类似的结果), 总有:

(1) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(l)}(\overrightarrow{\theta}:x,y)p^{(k)}(\overrightarrow{\theta};y,x)>0\})=1$ 蕴涵了 $d_0(i)|(l+k)$;

(2) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(n_1)}(\overrightarrow{\theta}:x,y)p^{(n_2)}(\overrightarrow{\theta};x,y)>0\})=1$ 蕴涵了 $d_0(i)|(n_1-n_2)$;

(3) 存在 $M(x)$, 使得当 $k\geqslant M(x)$ 时有:

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(kd_0(i))}(\overrightarrow{\theta};x,x)>0\})=1; \tag{4.19}$$

(4) 存在 $r_y$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(n)}(\overrightarrow{\theta};x,y)>0\})=1$$

蕴涵了 $n=r_y(\text{mod } d_0(i))$;

(5) 存在 $K(x)$, 当 $n\geqslant K(x)$ 时有:

$$\pi(\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(nd_0(i)+r_y)}(\overrightarrow{\theta};x,y)>0\})=1. \tag{4.20}$$

**证** (1) 设 (1) 中的条件成立, 由 (1.17) 式及 $\pi=\pi\circ T^{-1}$ 得

$$\begin{aligned}p^{(l+k)}(\overrightarrow{\theta};x,x)&\geqslant p^{(l)}(\overrightarrow{\theta};x,y)p^{(k)}(T^{l}\overrightarrow{\theta};y,x)\\&>0,\quad\pi-\text{a.s.}.\end{aligned} \tag{4.21}$$

所以 $d_0(i)|(l+k)$.

(2) 由于 $x,y\in\mathscr{R}_0(i)$, 仿 (4.2) 式, 存在正整数 $v$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(v)}(\overrightarrow{\theta};y,x)>0\})=1. \tag{4.22}$$

由 (1.17) 及 $\pi=\pi\circ T^{-1}$ 知

$$\begin{aligned} p^{(n_s+v)}(\overrightarrow{\theta};x,x) &\geqslant p^{(n_s)}(\overrightarrow{\theta};x,y)p^{(v)}(T^{n_s}\overrightarrow{\theta};y,x) \\ &> 0, \quad \pi-\text{a.s.} \quad (s=1,2). \end{aligned} \tag{4.23}$$

所以由 $d_0(i)(=d_x)$ 的定义知

$$d_0(i)|(n_1+v), \quad d_0(i)|(n_2+v),$$

从而 $d_0(i)|(n_1-n_2)$.

(3) 由 $d_0(i)(=d_x)$ 的定义得知: 存在 $n_1,n_2,\cdots,n_u$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(n_s)}(\overrightarrow{\theta};x,x)>0\})=1 \quad (1\leqslant s\leqslant u), \tag{4.24}$$

且

$$d_0(i)=\text{G.C.D.} \quad \{n_1,n_2,\cdots,n_u\}.$$

由初等数论的一条熟知的定理可知: 存在 $M(x)$, 对任何正整数 $k\geqslant M(x)$ 总有

$$kd_0(i)=\sum_{s=1}^{u}C_sn_s, \tag{4.25}$$

其中 $C_s$ 是依赖 $k$ 的非负整数. 因此, 由 $\pi=\pi\circ T^{-1}$ 及 (4.24) 和 (1.17) 得

$$\begin{aligned} p^{(kd_0(i))}(\overrightarrow{\theta};x,x) &= p^{(\sum\limits_{s=1}^{u}C_sn_s)}(\overrightarrow{\theta};x,x) \\ &= \prod_{s=1}^{u}p^{(C_sn_s)}\left(T^{\sum\limits_{r=1}^{s-1}C_rn_r}\overrightarrow{\theta};x,x\right)>0, \quad \pi-\text{a.s.}. \end{aligned} \tag{4.26}$$

(4) 因为 $x,y\in\mathscr{R}_0(i)$, 因此存在 $n'$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(n')}(\overrightarrow{\theta};x,y)>0\})=1.$$

取 $r_y=n'$, 则由 (2) 知 $d_0(i)|(n-r_y)$, 即 $r_y=n(\text{mod } d_0(i))$.

(5) 因为 $x,y\in\mathscr{R}_0(i)$, 所以

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:f^*_{\min}(\overrightarrow{\theta};x,y)>0\})=1,$$

再由 (4), 存在 $(k'd_0(i)+r_y)$ 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(k'd_0(i)+r_y)}(\overrightarrow{\theta};x,y)>0\})=1. \tag{4.27}$$

取 $K(x)=k'+M(x)$, 则当 $n\geqslant K(x)$ 时, 由 (3) 及 (1.17) 和 (4.27) 可证:

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(nd_0(i)+r_y)}(\overrightarrow{\theta};x,y)>0\})=1.$$

命题 4.2 证毕.

**命题 4.3** 对任何 $x,y\in\mathscr{R}_0(i)$ (当 $x,y\in\mathscr{R}_+(j)$ 亦有类似结果), 若 $x\approx y$, 则 $d_0(i)(=d_x)=H_{x,y}$.

**证** 由 $x,y\in\mathscr{R}_0(i),x\approx y$, 可取正整数 $t$, 使

$$\pi(\{\overrightarrow{\theta};p^{(t)}(\overrightarrow{\theta};x,y)>0\})=1. \tag{4.28}$$

由命题 4.2 (3), 我们可以取 $k_0$, 使

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(k_0d_0(i))}(\overrightarrow{\theta};x,x)p^{((k_0+1)d_0(i))}(\overrightarrow{\theta};x,x)>0\})=1. \tag{4.29}$$

因此, 由 (1.17)、(4.28)、(4.29) 可证:

$$\pi(\{\overrightarrow{\theta}:p^{(k_0d_0(i)+t)}(\overrightarrow{\theta};x,y)>0\})=1,$$
$$\pi(\{\overrightarrow{\theta}:p^{((k_0+1)d_0(i)+t)}(\overrightarrow{\theta};x,y)>0\})=1,$$

所以由 $H_{x,y}$ 之定义和上述两式知 $H_{x,y}|d_0(i)$. 而 $x\approx y$ 蕴涵了 $d_0(i)|H_{x,y}$. 总之 $H_{x,y}=d_0(i)$. 命题 4.3 证毕.

**命题 4.4** 设 $x,y,z\in\mathscr{R}_0(i)$ ($x,y,z\in\mathscr{R}_+(j)$ 亦有类似结果), 若

$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(k_1)}(\overrightarrow{\theta};x,y)>0\})=1,$$
$$\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}:p^{(k_2)}(\overrightarrow{\theta};x,z)>0\})=1,$$

而且 $k_1=k_2(\mathrm{mod}\ d_0(i))$, 则 $y\approx z$.

证明甚易, 读者可作为习题验证之.

**命题 4.5** $\mathscr{R}_0(i)$ 和 $\mathscr{R}_+(j)$ 可以根据等价关系 $\approx$ 分成一些互不相交的等价子类如下:

$$\mathscr{R}_0(i)=\sum_{s=1}^{d_0(i)}\mathscr{R}_0(i,s);$$
$$\mathscr{R}_+(j)=\sum_{t=1}^{d_+(j)}\mathscr{R}_0(j,t),$$

而且 $x$ 和 $y$ 同属于一个等价子类 $\mathscr{R}_0(i,s)$ (或 $\mathscr{R}_+(j)$) 的充分必要条件是 $x\approx y$.

证明甚易, 读者可作为习题验证之.

**定理 4.3** 应用前述符号, 状态空间 $\mathscr{X}$ 总可分解如下:

$$\begin{aligned}\mathscr{X} &= \mathscr{N}\bigcup\mathscr{R} = \mathscr{N}\bigcup(\mathscr{R}_0\bigcup\mathscr{R}_+)\\&= \mathscr{N}\bigcup\left(\bigcup_{i\in\Gamma_0}\mathscr{R}_0(i)\right)\bigcup\left(\bigcup_{j\in\Gamma_1}\mathscr{R}_+(j)\right)\\&= \mathscr{N}\bigcup\left(\bigcup_{i\in\Gamma_0}\bigcup_{s=1}^{d_0(i)}\mathscr{R}_0(i,s)\right)\bigcup\left(\bigcup_{j\in\Gamma_1}\bigcup_{t=1}^{d_+(j)}\mathscr{R}_+(j,t)\right)\end{aligned}$$

此外, 还有:

(1) 对任何 $i\in\Gamma_0, j\in\Gamma_1, \mathscr{R}_0(i)$ 和 $\mathscr{R}_+(j)$ 都是 $M$ 封闭的;

(2) $x$ 和 $y$ 属于同一个 $\mathscr{R}_0(i)$ (或 $\mathscr{R}_+(j)$) 的充分必要条件是: $x \overset{\min}{\longleftrightarrow} y$.

(3) $x$ 和 $y$ 属于同一个等价子集 $\mathscr{R}_0(i,s)$ (或 $\mathscr{R}_+(j,t)$) 的充分必要条件是: $x\approx y$.

(4) 对任何 $x,y\in\mathscr{R}_0(i)$ (或 $\mathscr{R}_+(j)$), 存在 $M(x), K(x)$ 和 $r_y$, 使

(i) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}: p^{(kd_x)}(\overrightarrow{\theta};x,x)>0\})=1 \quad (\forall k\geqslant M(x))$;

(ii) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}: p^{(n)}(\overrightarrow{\theta};x,y)>0\})=1$ 蕴涵了 $n=r_y \pmod{d_x}$;

(iii) $\pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}: p^{(nd_x+r_y)}(\overrightarrow{\theta};x,y)>0\})=1 \quad (\forall n\geqslant K(x))$.

(5) $x\in\mathscr{R}_0(i,r), y\in\mathscr{R}_0(i,r+1) \Rightarrow \pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}: p^{(1)}(\overrightarrow{\theta};x,y)=0\})=1$; $x\in\mathscr{R}_+(j,r), y\in\mathscr{R}_+(j,r+1) \Rightarrow \pi(\{\overrightarrow{\theta}\in\overrightarrow{\Theta}: p^{(1)}(\overrightarrow{\theta};x,y)=0\})=1$.

**证** 本定理是前面一系结论的总结. 读者可作为小结一一验证之.

## §5 依时随机环境中的分枝链

如不特别声明, 本节仍沿袭前四节的符号. 本节特别地取 $\mathscr{X}=\{0,1,2,\cdots\}$, $\mathscr{A}$ 仍为 $\mathscr{X}$ 中一切子集所成的 $\sigma$ 代数, $(\Theta,\mathscr{B})$ 是任一可测空间, $M(\mathscr{X},\mathscr{A})$ 表示 $\mathscr{X}\times\mathscr{X}$ 上的一切转移矩阵, $RM(\Theta;\mathscr{X},\mathscr{A})$ 表一切随机转移矩阵. 对于任何 $p(\cdot)\in RM(\Theta;\mathscr{X},\mathscr{A}), p^{(n)}(\overrightarrow{\theta};x,y)$ 的定义如 (1.2), $p^{(n)}(\overrightarrow{\theta})$ 的定义如 (1.1). 为了与本节的微商的符号混淆, 记 $p^{(n)}(\overrightarrow{\theta})$、$p^{(n)}(\overrightarrow{\theta};x,y)$ 分别为 $p_n(\overrightarrow{\theta})$ 和 $p_n(\overrightarrow{\theta};x,y)$.

**定义 5.1** 称概率空间 $(\Omega,\mathscr{F},P)$ 上的具有随机转移矩阵 $p(\cdot)$ 的 MCTRE $(\overrightarrow{X},\overrightarrow{\xi})$ 是一个**依时随机环境中的分枝链** (简记为 BCTRE), 如果 $p(\cdot)$ 满足:

$$(\text{MB}): p(\theta;i,j)=\sum_{\sum\limits_{s=1}^{i} j_s=j}\prod_{s=1}^{i}p(\theta;1,j_s) \tag{5.1}$$

对一切 $\theta\in\Theta, i,j\geqslant 0$ 成立. 当 $i=0$ 时, (5.1) 式右方定义为 $\delta_{i,j}$.

本节恒设 $X_0 \equiv a, a$ 是正整数.

**定义 5.2** 对任何具有随机转移矩阵 $p(\bullet)$ 的 BCTRE $(\overrightarrow{X}, \overrightarrow{\xi})$ 来说, 称

$$\varphi(\theta; s) \stackrel{\text{def.}}{=} \sum_{k=0}^{\infty} p(\theta; 1, k) s^k, \quad (\theta \in \Theta, |s| \leqslant 1), \tag{5.2}$$

为**分枝母函数**, 而称 $\{p(\theta; 1, k), k \geqslant 0\}$ 为**分枝律**. 简记 $\varphi(\theta; \bullet) \stackrel{\text{def.}}{=} \varphi(\theta)$. 称

$$\varphi_n(\overrightarrow{\theta}; s) \stackrel{\text{def.}}{=} \sum_{k=0}^{\infty} p_n(\overrightarrow{\theta}; 1, k) s^k, \quad (n \geqslant 0, \overrightarrow{\theta} \in \overrightarrow{\Theta}, |s| \leqslant 1) \tag{5.3}$$

为 $n$ 次分枝母函数. 简记 $\varphi_n(\overrightarrow{\theta}; \bullet) \stackrel{\text{def.}}{=} \varphi_n(\overrightarrow{\theta})$. 称

$$\psi_n(\overrightarrow{\theta}; s) \stackrel{\text{def.}}{=} \sum_{k=0}^{\infty} P(X_n = k | \overrightarrow{\xi} = \overrightarrow{\theta}) s^k \quad (n \geqslant 0, \overrightarrow{\theta} \in \overrightarrow{\Theta}, |s| \leqslant 1), \tag{5.4}$$

为在环境 $\overrightarrow{\xi} = \overrightarrow{\theta}$ 下, $X_n$ 的母函数.

记 $\varphi_n^{(k)}(\overrightarrow{\theta}; s)$ 和 $\psi_n^{(k)}(\overrightarrow{\theta}; s)$ 分别为 $\varphi_n(\overrightarrow{\theta}; s)$ 和 $\psi_n(\overrightarrow{\theta}; s)$ 对 $s$ 的 $k$ 阶微商 $(|s| < 1)$.

$$\varphi_n^{(k)}(\overrightarrow{\theta}; 1) \stackrel{\text{def.}}{=} \lim_{s \uparrow 1} \varphi_n^{(k)}(\overrightarrow{\theta}; s), \quad \psi_n^{(k)}(\overrightarrow{\theta}; 1) \stackrel{\text{def.}}{=} \lim_{s \uparrow 1} \psi_n^{(k)}(\overrightarrow{\theta}; s).$$

**命题 5.1** 设 $(\overrightarrow{X}, \overrightarrow{\xi})$ 是具有随机转移矩阵 $p(\bullet)$ 的 BCTRE, $\varphi, \varphi_n$ 和 $\psi_n$ 如 (5.2)、(5.3) 和 (5.4) 所定义, 则

(1) $P(X_n = k | \overrightarrow{\xi} = \overrightarrow{\theta}) = p_n(\overrightarrow{\theta}; a, k);$ (5.5)

(2) $\varphi_n(\overrightarrow{\theta}; s) = \varphi_{n-1}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s))$ (5.6)

$$(\forall n \geqslant 1, \overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}, |s| \leqslant 1);$$

(3) $\psi_n(\overrightarrow{\theta}) = [\varphi(\theta_0) \circ \varphi(\theta_1) \circ \cdots \circ \varphi(\theta_{n-1})]^a$ (5.7)

$$(\forall n \geqslant 1, \overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}).$$

特别地, $\psi_n(\overrightarrow{\theta}; s) = \psi_{n-1}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s))$,

$$\psi_1(\overrightarrow{\theta}) = \varphi(\theta_0)^a, \quad \psi_0(\overrightarrow{\theta}; s) = s^a.$$

**证** (1) 由 $X_0 \equiv a$ 及 $p_n(\overrightarrow{\theta}; i, j)$ 的定义即得.

(2) 由于 $p(\bullet)$ 满足 (5.1), 而且

$$p_1(\overrightarrow{\theta}; j, k) = p(\theta_0; j, k),$$

所以

$$\begin{aligned}\sum_{k=0}^{\infty} p_1(\overrightarrow{\theta}; j, k)s^k &= \sum_{k=0}^{\infty} p(\theta_0; j, k)s^k \\ &\xlongequal{(5.1)} \left(\sum_{k=0}^{\infty} p(\theta_0; 1, k)s^k\right)^j = \varphi(\theta_0; s)^j.\end{aligned} \tag{5.8}$$

再由 (1.17),

$$\begin{aligned}\varphi_n(\overrightarrow{\theta}; s) &= \sum_{k=0}^{\infty} p_n(\overrightarrow{\theta}; 1, k)s^k \\ &\xlongequal{(1.17)} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} p_{n-1}(\overrightarrow{\theta}; 1, j)p_1(T^{n-1}\overrightarrow{\theta}; j, k)s^k \\ &\xlongequal{(5.8)} \sum_{j=0}^{\infty} p_{n-1}(\overrightarrow{\theta}; 1, j)\varphi(\theta_{n-1}; s)^j \\ &= \varphi_{n-1}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s)).\end{aligned}$$

(3)

$$\begin{aligned}\psi_n(\overrightarrow{\theta}; s) &= \sum_{k=0}^{\infty} P(X_n = k | \overrightarrow{\xi} = \overrightarrow{\theta})s^k \\ &= \sum_{k=0}^{\infty} p_n(\overrightarrow{\theta}; a, k)s^k \\ &\xlongequal{(1.17)} \sum_{j=0}^{\infty} p_{n-1}(\overrightarrow{\theta}; a, j)\sum_{k=0}^{\infty} p_1(T^{n-1}\overrightarrow{\theta}; j, k)s^k \\ &\xlongequal{(5.1)} \sum_{j=0}^{\infty} p_{n-1}(\overrightarrow{\theta}; a, j)\varphi(\theta_{n-1}; s)^j \\ &= \psi_{n-1}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s)).\end{aligned}$$

仿此, 继续作下去, 并注意 $\psi_0(\overrightarrow{\theta}; s) = s^a$ 即可证 (3) 成立.

**命题 5.2**　设 $(\overrightarrow{X}, \overrightarrow{\xi})$ 是具有随机转移矩阵 $p(\cdot)$ 的 BCTRE, $\varphi, \varphi_n$ 和 $\psi_n$ 如 (5.2)、(5.3) 和 (5.4) 所定义, 则

(1)
$$\boldsymbol{E}(X_n | \overrightarrow{\xi} = \overrightarrow{\theta}) = a\prod_{i=0}^{n-1} b_i \quad (n \geqslant 1); \tag{5.9}$$

(2)
$$\begin{aligned}&\operatorname{var}(X_n | \overrightarrow{\xi} = \overrightarrow{\theta}) \\ &= a\left[\sum_{i=0}^{n-1} \frac{a_i}{b_i} \prod_{k=i+1}^{n-1} b_i + 1 - \prod_{i=0}^{n-1} b_i\right]\prod_{i=0}^{n-1} b_i,\end{aligned} \tag{5.10}$$

其中

$a_n = \varphi^{(2)}(\theta_n; 1), b_n = \varphi^{(1)}(\theta_n; 1), \overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}).$

**证** 令

$$\alpha_n = \varphi_n^{(2)}(\overrightarrow{\theta}; 1), \quad \beta_n = \varphi_n^{(1)}(\overrightarrow{\theta}; 1).$$

(1) 由 (5.4) 有

$$\begin{aligned} \boldsymbol{E}(X_n | \overrightarrow{\xi} = \overrightarrow{\theta}) &= \psi_n^{(1)}(\overrightarrow{\theta}; 1) \\ &= a \prod_{i=0}^{n-1} \varphi^{(1)}(\theta_i; 1) = a \prod_{i=0}^{n-1} b_i. \end{aligned}$$

(2) 由命题 5.1 (2) 有

$$\varphi_n(\overrightarrow{\theta}; s) = \varphi_{n-1}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s)).$$

所以

$$\varphi_n^{(1)}(\overrightarrow{\theta}; s) = \varphi_{n-1}^{(1)}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s))\varphi^{(1)}(\theta_{n-1}; s). \tag{5.11}$$

从而

$$\begin{aligned} \varphi_n^{(2)}(\overrightarrow{\theta}; s) = {} & \varphi_{n-1}^{(2)}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s)) \left[\varphi^{(1)}(\theta_{n-1}; s)\right]^2 \\ & + \varphi_{n-1}^{(1)}(\overrightarrow{\theta}; \varphi(\theta_{n-1}; s))\varphi^{(2)}(\theta_{n-1}; s). \end{aligned}$$

对 $n$ 作归纳法并注意: $\beta_{n+1} = \beta_n b_n, \beta_0 = 1, \alpha_1 = a_0$ 可得

$$\begin{aligned} \alpha_n &= \varphi_n^{(2)}(\overrightarrow{\theta}; 1) \\ &= \alpha_{n-1} b_{n-1}^2 + \beta_{n-1} a_{n-1} = \cdots \\ &= \alpha_1 \prod_{i=1}^{n-1} b_i^2 + \sum_{k=1}^{n-1} a_k \beta_k \prod_{i=k+1}^{n-1} b_i^2 \\ &= a_0 \prod_{i=1}^{n-1} b_i^2 + \sum_{k=1}^{n-1} a_k \left(\prod_{t=0}^{k-1} b_t\right) \prod_{i=k+1}^{n-1} b_i^2 \\ &= \left(\prod_{i=0}^{n-1} b_i\right) \left[\frac{a_0}{b_0} \prod_{i=1}^{n-1} b_i + \sum_{k=1}^{n-1} \frac{a_k}{b_k} \prod_{i=k+1}^{n-1} b_i\right] \\ &= \left(\prod_{i=0}^{n-1} b_i\right) \sum_{i=0}^{n-1} \frac{a_i}{b_i} \prod_{k=i+1}^{n-1} b_k. \end{aligned} \tag{5.12}$$

但

$$\psi_n^{(1)}(\overrightarrow{\theta}; 1) = a\beta_n, \tag{5.13}$$

$$\psi_n^{(2)}(\overrightarrow{\theta}; 1) = a(a-1)\beta_n^2 + a\alpha_n, \tag{5.14}$$

由 (5.12)、(5.13)、(5.14) 得

$$\begin{aligned}
&\operatorname{var}(X_n|\overrightarrow{\xi}=\overrightarrow{\theta})\\
&=\psi_n^{(2)}(\overrightarrow{\theta};1)+\psi_n^{(1)}(\overrightarrow{\theta};1)-(\psi_n^{(1)}(\overrightarrow{\theta};1))^2\\
&=a(a-1)\beta_n^2+a\alpha_n+a\beta_n-a^2\beta_n^2\\
&=a(\alpha_n+\beta_n-\beta_n^2)\\
&=a\left[\sum_{i=0}^{n-1}\frac{a_i}{b_i}\prod_{k=i+1}^{n-1}b_k+1-\prod_{i=0}^{n-1}b_i\right]\prod_{i=0}^{n-1}b_i.
\end{aligned}$$

命题证毕.

下面我们研究 BCTRE 的构造.

**定义 5.3** 称 $\{\rho_j(\theta),j\geqslant 0\}$ 是一个**随机分布**, 如果 $\rho_j(\theta)\geqslant 0,\sum\limits_{j\geqslant 0}\rho_j(\theta)=1\ (\forall j\geqslant 0,\theta\in\Theta)$, 而且对任意固定的 $j\geqslant 0,\rho_j(\bullet)$ 是 $\mathscr{B}$ 可测函数.

**定义 5.4** 称 MCTRE $(\overrightarrow{X},\overrightarrow{\xi})$ 是一个**典范的依时随机环境中的分枝链**, (简记为 CBCTRE), 如果存在一族关于 $\overrightarrow{\xi}$ 条件独立的随机变量 $\{Y_k^{(n)},n\geqslant 1,k\geqslant 1\}$, 即是

$$P\left(\bigcap_{n=1}^{N}\bigcap_{k=1}^{K}\{Y_k^{(n)}=y_k^{(n)}\}|\overrightarrow{\xi}\right)=\prod_{n=1}^{N}\prod_{k=1}^{K}P(\{Y_k^{(n)}=y_k^{(n)}\}|\overrightarrow{\xi}),\tag{5.15}$$

对任何 $N\geqslant 1,K\geqslant 1$, 任何非负整数族 $\{y_k^{(n)},1\leqslant n\leqslant N,1\leqslant k\leqslant K\}$ 成立, 此外, 对任何 $n\geqslant 1,y\geqslant 0,P(Y_k^{(n)}=y|\overrightarrow{\xi})$ 不依赖 $k\geqslant 1$. $\{Y_k^{(n)}\}$ 使

$$X_{n+1}=\sum_{k=1}^{X_n}Y_k^{(n+1)}(n\geqslant 0),\quad X_0\equiv a,\tag{5.16}$$

其中 $a$ 是正整数.

**定理 5.1** 任给一个随机分布 $\{\rho_k(\theta),k\geqslant 0\}$ 和 $\overrightarrow{\mathscr{B}}$ 上的一个概率测度 $\pi$, 恒存在一个概率空间 $(\Omega,\mathscr{F},P)$ 及定义在其上的一个 CBCTRE $\{\overrightarrow{X},\overrightarrow{\xi}\}$ 使

(1) $(\overrightarrow{X},\overrightarrow{\xi})$ 的分枝律为 $\{\rho_k(\theta),k\geqslant 0\}$, 其分枝母函数

$$\varphi(\theta;s)=\sum_{k=0}^{\infty}\rho_k(\theta)s^k;\tag{5.17}$$

(2) $\overrightarrow{\xi}$ 的 $P$ 边缘分布 $P\circ(\overrightarrow{\xi})^{-1}=\pi$;

(3) 生成随机变量 $Y_k^{(n+1)}$ 的条件母函数为

$$\varphi(\theta_n; s) = \sum_{j \geqslant 0} P(Y_k^{n+1} = j | \overrightarrow{\xi} = \overrightarrow{\theta}) \tag{5.18}$$

$(\forall \overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}, |s| \leqslant 1, n \geqslant 0, k \geqslant 1)$, 其中 $\varphi(\theta; s)$ 由 (5.17) 式所定义.

**证** 记

$M(\mathscr{X}) \stackrel{\text{def.}}{=\!=} \{A = (A_{i,j}, i, j \geqslant 0) : A_{i,j} \in \mathscr{X}\}$ 为一切元均在 $\mathscr{X}$ 中全体矩阵,

$$\begin{aligned}
&\Omega \stackrel{\text{def.}}{=\!=} M(\mathscr{X}) \times \overrightarrow{\Theta}, \\
&\mathscr{F} \stackrel{\text{def.}}{=\!=} (\mathscr{A}^{\mathbf{N}_+ \times \mathbf{N}_+}) \times \overrightarrow{\mathscr{B}}, \\
&L(M(\mathscr{X}) : a_{i,j}, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant k) \\
\stackrel{\text{def.}}{=\!=} &\{A \in M(\mathscr{X}) : A_{i,j} = a_{i,j}, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant k\}
\end{aligned}$$

为 $\mathscr{A}^{\mathbf{N}_+ \times \mathbf{N}_+}$ 中的可测柱集, 此处 $\mathbf{N}_+ = \{1, 2, \cdots\}, a_{i,j} \in \mathscr{X}$.

先在可测柱集上定义集函数 $P_{\overrightarrow{\theta}}$ 如下:

$$\begin{aligned}
&P_{\overrightarrow{\theta}}(L(M(\mathscr{X}) : a_{i,j}, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant k)) \\
= &\prod_{i=1}^{n} \prod_{j=1}^{k} \rho_{a_{i,j}}(\theta_{i-1}), \\
&(\overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}, a_{i,j} \in \mathscr{X}, n \geqslant 1, k \geqslant 1).
\end{aligned} \tag{5.19}$$

易证

$$\sum_{i=1}^{n} \sum_{j=1}^{k} \left( \sum_{a_{i,j} \in \mathscr{X}} \prod_{i=1}^{n} \prod_{j=1}^{k} \rho_{a_{i,j}}(\theta_{i-1}) \right) = 1. \tag{5.20}$$

此定义在可测柱集上的集函数可唯一地扩张到 $\mathscr{A}^{\mathbf{N}_+ \times \mathbf{N}_+}$ 上去而得一概率测度, 此概率测度仍用 $P_{\overrightarrow{\theta}}$ 记之, 此外, 对任何固定的 $D \in \mathscr{A}^{\mathbf{N}_+ \times \mathbf{N}_+}, P_{\overrightarrow{\theta}}(D)$ 作为 $\overrightarrow{\theta}$ 的函数是 $\overrightarrow{\mathscr{B}}$ 可测的.

定义

$$P(C) = \int_{\overrightarrow{\Theta}} \pi(d\overrightarrow{\theta}) \int_{M(\mathscr{X})} P_{\overrightarrow{\theta}}(dA) \mathbf{1}_C(A, \overrightarrow{\theta}), \tag{5.21}$$

此处 $C \in \mathscr{F}$. 显然 $P$ 是 $\mathscr{F}$ 上的一个概率测度, 于是我们得一概率空间 $(\Omega, \mathscr{F}, P)$.

对任一 $\omega=(A,\vec{\theta})\in\Omega=M(\mathscr{X})\times\vec{\Theta}, A=(A_{i,j},i,j\geqslant 0)$, 定义

$$\vec{\xi}(\omega)=\vec{\theta}, Y_j^{(i)}(\omega)=A_{i,j},\ (i,j\geqslant 1), \tag{5.22}$$

$$X_{n+1}(\omega)=\sum_{k=1}^{X_n(\omega)} Y_k^{(n+1)}(\omega)(n\geqslant 0), X_0(\omega)=a, \tag{5.23}$$

$$p(\theta;x,y)=\sum_{\sum_{k=1}^{x} y_k=y}\prod_{k=1}^{x}\rho_{y_k}(\theta), \theta\in\Theta, x,y\in\mathscr{X}. \tag{5.24}$$

要证 $(\vec{X},\vec{\xi})$ 是一个具有随机转移矩阵 $p(\cdot)$ 的 CBCRE $(p(\theta)(x,y)\stackrel{\text{def.}}{=}p(\theta;x,y))$, 且满足定理 5.1 中的 (1)、(2)、(3).

由 (5.21) 易见: 对任何 $B\in\vec{\mathscr{B}}$, 有

$$\begin{aligned}&P\circ(\vec{\xi})^{-1}(B)=P(M(\mathscr{X})\times B)\\&=\int_{\vec{\Theta}}\pi(d\vec{\theta})P_{\vec{\theta}}(M(\mathscr{X}))1_B(\vec{\theta})=\pi(B).\end{aligned} \tag{5.25}$$

此即结论 (2) 成立.

对于任何 $a_{i,j}\in\mathscr{X}, B\in\vec{\mathscr{B}}$, 由 (5.21) 和 (5.25) 知

$$\begin{aligned}&P\left(\bigcap_{i=1}^{n}\bigcap_{j=1}^{k}\{Y_j^{(i)}=a_{i,j}\},\vec{\xi}\in B\right)\\&=P(L(M(\mathscr{X}):a_{i,j},1\leqslant i\leqslant n,1\leqslant j\leqslant k)\times B)\\&=\int_B\pi(d\vec{\theta})P_{\vec{\theta}}(L(M(\mathscr{X}):a_{i,j},1\leqslant i\leqslant n,1\leqslant j\leqslant k))\\&=\int_B P\circ(\vec{\xi})^{-1}(d\vec{\theta})\prod_{i=1}^{n}\prod_{j=1}^{k}\rho_{a_{i,j}}(\theta_{i-1}),\end{aligned} \tag{5.26}$$

因此

$$\begin{aligned}&P\left(\bigcap_{i=1}^{n}\bigcap_{j=1}^{k}\{Y_j^{(i)}=a_{i,j}\}|\vec{\xi}=\vec{\theta}\right)=\prod_{i=1}^{n}\prod_{j=1}^{k}\rho_{a_{i,j}}(\theta_{i-1})\\&=\prod_{i=1}^{n}\prod_{j=1}^{k}P(Y_j^{(i)}=a_{i,j}|\vec{\xi}=\vec{\theta})\quad(n,k\geqslant 1,a_{i,j}\in\mathscr{X}).\end{aligned} \tag{5.27}$$

(5.27) 说明 $\{Y_j^{(i)},i\geqslant 1,j\geqslant 1\}$ 关于 $\vec{\xi}$ 是条件独立的, 而且 $Y_j^{(i)}$ 的条件母函数为

$$\begin{aligned}&\sum_{k=0}^{\infty}P(Y_j^{(i)}=k|\vec{\xi}=\vec{\theta})s^k=\sum_{k=0}^{\infty}\rho_k(\theta_{i-1})s^k\\&=\varphi(\theta_{i-1};s)\quad(|s|\leqslant 1).\end{aligned} \tag{5.28}$$

(5.28) 说明结论 (3) 成立.

由 (5.23) 知 $p(\theta;1,j)=\rho_j(\theta)$. 结论 (1) 成立.

最后我们证明 $(\vec{X},\vec{\xi})$ 是具有随机转移矩阵 $p(\cdot)$ 的 CBCTRE. 事实上, 由于 $\{Y_k^{(n)},n\geqslant 1,k\geqslant 1\}$ 关于 $\vec{\xi}$ 条件独立, 若记 $P(\cdot|\vec{\xi})$ 为 $\widehat{P}_{\vec{\xi}}$, 则由 (5.23)、(5.24) 知

$$\begin{aligned}
&\widehat{P}_{\vec{\xi}}(X_{n+1}=y|X_0=x_0,\cdots,X_{n-1}=x_{n-1},X_n=x)\\
&=\widehat{P}_{\vec{\xi}}\left(\sum_{k=1}^{x}Y_k^{(n+1)}=y|X_n=x\right)\\
&=\widehat{P}_{\vec{\xi}}\left(\sum_{k=1}^{x}Y_k^{(n+1)}=y\right)\\
&=\sum_{\sum\limits_{k=1}^{x}y_k=y}\prod_{k=1}^{x}P_{(\xi)}(Y_k^{(n+1)}=y_k)\\
&=\sum_{\sum\limits_{k=1}^{x}y_k=y}\prod_{k=1}^{x}\rho_{y_k}(\xi_n)\\
&=p(\xi_n;x,y)\quad(n\geqslant 0,x,y\in\mathscr{X}).
\end{aligned}\tag{5.29}$$

显然

$$P(X_0=x|\vec{\xi})=\delta_{a,x}=P(X_0=x|\vec{\xi}^{\,0}_{-\infty}),\tag{5.30}$$

由 (5.29)、(5.30) 及 (5.23) 和定理 1.2 得知 $(\vec{X},\vec{\xi})$ 是 CBCTRE, 具有随机转移矩阵 $p(\cdot)$, 定理证毕.

**定理 5.2** 设 $\vec{X}=\{X_0,X_1,\cdots\}$、$\vec{\xi}=\{\xi_n,n\in\mathbf{Z}\},Y=\{Y_k^{(n)},n\geqslant 1,k\geqslant 1\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的三族随机变量, $X_m$ 和 $Y_k^{(n)}$ 取值于 $\mathscr{X}$, $(m\geqslant 0,n,k\geqslant 1)$, $\xi_n$ 取值于 $\Theta$, $\{\rho_j(\theta),j\geqslant 0\}$ 是一个随机分布, $\varphi(\theta;s)=\sum\limits_{j\geqslant 0}\rho_j(\theta)s^j$ $(|s|\leqslant 1)$ 是其母函数, 如果 $Y_k^{(n)}$ 关于 $\vec{\xi}$ 的条件母函数是:

$$\begin{aligned}
&\sum_{j=0}^{\infty}P(Y_k^{(n)}=j|\vec{\xi})s^j=\varphi(\xi_{n-1};s)\\
&(|s|\leqslant 1,n\geqslant 1,k\geqslant 1),
\end{aligned}$$

$X_0\equiv a$ 是正整数, $X_{n+1}=\sum\limits_{k=1}^{X_n}Y_k^{(n+1)}$ $(n\geqslant 0)$, 则 $(\vec{X},\vec{\xi})$ 是一个具有下述

随机转移矩阵 $p(\cdot)$ 的 MCTRE, 从而是一个 CBCTRE:

$$p(\theta;i,j)=\sum_{\sum_{s=1}^{i}j_s=j}\prod_{s=1}^{i}\rho_{j_s}(\theta).$$

证明甚易, 读者可作为习题验证之.

**定理 5.3** (1) CBCTRE 一定是 BCTRE.

(2) 如果 $(\overrightarrow{X},\overrightarrow{\xi})$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的, 具有随机转移矩阵 $p(\cdot)$ 的 BCTRE, $X_n$ 和 $\xi_n$ 分别取值于 $\mathscr{X}$ 和 $\Theta$, 则必存在一个定义在概率空间 $(\Omega^*,\mathscr{F}^*,P^*)$ 上的 CBCTRE $(\overrightarrow{X}^*,\overrightarrow{\xi}^*)$, $X_n^*$ 和 $\xi_n^*$ 分别取值于 $\mathscr{X}$ 和 $\Theta$, 而且 $(\overrightarrow{X},\overrightarrow{\xi})$ 的 $P$ 分布 $P\circ(\overrightarrow{X},\overrightarrow{\xi})^{-1}$ 和 $(\overrightarrow{X}^*,\overrightarrow{\xi}^*)$ 的 $P^*$ 分布 $P^*\circ(\overrightarrow{X}^*,\overrightarrow{\xi}^*)^{-1}$ 一样.

证明甚易, 读者可作为习题验证之.

## §6 依时且依空随机环境中的 Markov 链简介

在这一节中, 我们将简单地介绍一下依时且依空的随机环境中的 Markov 链. 随机环境是一族既依时又依空的随机变量, 其状态空间是任一可测空间, 而且本原链的状态空间也不一定可数.

设 $(\mathscr{X},\mathscr{A})$ 和 $(\Gamma,\mathscr{D})$ 是两个任意的可测空间, $\Theta$ 是一族由 $\mathscr{X}$ 到 $\Gamma$ 的 $\mathscr{A}/\mathscr{D}$ 可测函数, $\mathscr{B}$ 是 $\Theta$ 上的一个 $\sigma$ 代数, 于是有第三个可测空间 $(\Theta,\mathscr{B})$. 仍令 $\mathbf{Z}=\{0,\pm1,\pm2\cdots\}$ 为整数集, $\mathbf{N}=\mathbf{Z}_+=\{0,1,2,\cdots\}$ 为非负整数集, $(\Omega,\mathscr{F},P)$ 是完备概率空间.

令 $\overrightarrow{X}=\{X_n,n\in\mathbf{N}\}$ 和 $\overrightarrow{\xi}=\{\xi_{m,x},m\in\mathbf{N},x\in\mathscr{X}\}$ 是 $\Omega$ 上的分别取值于 $\mathscr{X}$ 和 $\Gamma$ 的两族随机变量, 且 $\xi_{m,\cdot}(\cdot):\mathscr{X}\times\Omega\mapsto\Gamma$ 是 $(\mathscr{A}\times\mathscr{F})/\mathscr{D}$ 可测的. 令 $\xi_n=\{\xi_{n,x},x\in\mathscr{X}\},(n\in\mathbf{N}),\xi_{(x)}=\{\xi_{n,x},n\in\mathbf{N}\}\quad(x\in\mathscr{X})$.

定义

$$\overrightarrow{X}_k^n=\{X_m,k-1<m<n+1\},\quad 0\leqslant k\leqslant n\leqslant\infty,$$
$$\overrightarrow{\xi}_k^n=\{\xi_m,k-1<m<n+1\},\quad -\infty\leqslant k\leqslant n\leqslant\infty.$$

注意 $\overrightarrow{X}=\overrightarrow{X}_0^\infty,\overrightarrow{\xi}=\overrightarrow{\xi}_{-\infty}^\infty$.

再令

$$\overrightarrow{\Theta}=\Theta^{\mathbf{Z}},\quad \overrightarrow{\mathscr{B}}=\mathscr{B}^{\mathbf{Z}},$$
$$\overrightarrow{\mathscr{B}}_k^n=\prod_{k-1<m<n+1}\mathscr{B}_m,\quad \overrightarrow{\Theta}_k^n=\prod_{k-1<m<n+1}\Theta_m,\quad -\infty\leqslant k\leqslant n\leqslant\infty,$$
$$\mathscr{B}_m\equiv\mathscr{B}=\Theta\bigcap\mathscr{D}^{\mathscr{X}},\quad \Theta_m\equiv\Theta.$$

**定义 6.1** 设 $p:\mathscr{X}\times\mathscr{A}\mapsto[0,1]$, 称 $p$ 是可测空间 $(\mathscr{X},\mathscr{A})$ 上的一个 Markov 核, 简称 Markov 核, 如果对任何固定的 $x\in\mathscr{X}, p(x,\bullet)$ 是 $\mathscr{A}$ 上的概率测度; 对任何固定的 $A\in\mathscr{A}, p(\bullet,A)$ 是 $\mathscr{A}$ 可测函数. 记全体 Markov 核为 $MK$.

设 $p(\gamma;\bullet,\bullet):\Gamma\mapsto MK$, 即 $\{p(\gamma),\gamma\in\Gamma\}$ 是一族 Markov 核, 若对任何固定的 $A\in\mathscr{A}, p(\bullet;\bullet,A)$ 是 $\mathscr{D}\times\mathscr{A}$ 可测函数, 则称 $p(\gamma;x,A)$ (或简记 $p$) 为环境 $\Gamma$ 中的一个随机 Markov 核, 简称为随机 Markov 核. 记全体 $\Gamma$ 中的随机 Markov 核为 $\mathrm{RMK}-\Gamma$, 简记为 RMK. 有时记 $p(\gamma;x,A)$ 为 $p(\gamma)$.

显然, 若 $p(x,A)$ 是 Markov 核, 则

$$p^{(n)}(x,A)\overset{\text{def.}}{=}\int_{\mathscr{X}}p(x,dx_1)\int_{\mathscr{X}}p(x_1,dx_2)\cdots\int_{\mathscr{X}}p(x_{n-1},dx_n)\mathbf{1}_A(x_n)$$

是转移函数 ($n\geqslant1, x_0$ 定义为 $x$).

**定义 6.2** 设 $(\Omega,\mathscr{F},P),(\Gamma,\mathscr{D}),(\mathscr{X},\mathscr{A}),(\Theta,\mathscr{B}),\vec{X},\vec{\xi},\vec{X}_k^n,\vec{\xi}_k^n$ 如前定义, $p(\gamma;x,A)\in\mathrm{RMK}-\Gamma$. 如下列条件 (M) 成立:

$$\begin{aligned}&(\mathrm{M}):P\left(X_{n+1}\in A,\xi_{n+1}\in B|\vec{X}_0^n,\vec{\xi}_{-\infty}^n\right)\\&=p(\xi_{n,X_n};X_n,A)P(\xi_{n+1}\in B|\vec{\xi}_{-\infty}^n),\end{aligned}\tag{6.1}$$

其中 $n\in\mathbf{N},A\in\mathscr{A},B\in\mathscr{B}$, 则称 $\{\vec{X},\vec{\xi}\}$ 是一个具有随机 Markov 核 $p(\gamma)\in\mathrm{RMK}-\Gamma$ 的**依时且依空的随机环境中的 Markov 链** (MCSTRE), $\vec{X}$ 称为本原链, $\vec{\xi}$ 称为环境随机场, 简称为环境.

如果 $(\vec{X},\vec{\xi}_0^\infty)$ 满足

$$\begin{aligned}&(\mathrm{M}^+):P(X_{n+1}\in A,\xi_{n+1}\in B|\vec{X}_0^n,\vec{\xi}_0^n)\\&=p(\xi_{n,X_n};X_n,A)P(\xi_{n+1}\in B|\vec{\xi}_0^n),\end{aligned}\tag{6.2}$$

其中 $A\in\mathscr{A},B\in\mathscr{B},n\in\mathbf{N}$, 则称 $\{\vec{X},\vec{\xi}_0^\infty\}$ 是一个具有随机 Markov 核 $p(\gamma)\in\mathrm{RMK}-\Gamma$ 的依正时且依空的随机环境中的 Markov 链 (MCSTRE(+)).

特别地, 若对每个 $n\in\mathbf{Z},\xi_{n,x}$ 都不依赖于 $x$, 则称 $\{\vec{X},\vec{\xi}\}$ 是一个依时的随机环境中的 Markov 链 (MCTRE); 类似地, 若对每个 $x\in\mathscr{X},\xi_{n,x}$ 都不依赖 $n$ 则称 $\{\vec{X},\vec{\xi}\}$ 是一个依空的随机环境中的 Markov 链 (MCSRE).

**附注 6.1** Cogburn [15] 意义下的随机环境中的 Markov 链, 就是定义 6.2 中的 MCTRE; Zeitouni [119] 意义下的随机环境中的 Markov 链, 就是定义 6.2 中的 MCSRE.

**附注 6.2** 对于任一 $p(\gamma)\in\mathrm{RMK}-\Gamma, p(\theta(x);x,A)$ 是 $x$ 的 $\mathscr{A}$ 可测函数 (对任何固定的 $A\in\mathscr{A}$ 和 $\theta\in\Theta$).

由单调类定理, 容易证明: 条件 (M) 等价于:

$$(\widehat{\mathrm{M}}): P(X_{n+1} \in A, \bigcap_{i=1}^{k}\{\xi_{n+1,x_i} \in D_i\}|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n)$$

$$= p(\xi_{n,X_n}; X_n, A)P\left(\bigcap_{i=1}^{k}\{\xi_{n+1,x_i} \in D_i\}|\overrightarrow{\xi}_{-\infty}^n\right), \quad (6.1)'$$

其中 $n \in \mathbf{N}, A \in \mathscr{A}, \{x_1, \cdots, x_k\} \subset \mathscr{X}, \{D_1, \cdots, D_k\} \subset \mathscr{D}, k \geqslant 1$;

条件 $(\mathrm{M}^+)$ 等价于

$$(\widehat{\mathrm{M}}^+): P(X_{n+1} \in A, \bigcap_{i=1}^{k}\{\xi_{n+1,x_i} \in D_i\}|\overrightarrow{X}_0^n, \overrightarrow{\xi}_0^n)$$

$$= p(\xi_{n,X_n}; X_n, A)P\left(\bigcap_{i=1}^{k}\{\xi_{n+1,x_i} \in D_i\}|\overrightarrow{\xi}_0^n\right), \quad (6.2$$

其中 $n \in \mathbf{N}, A \in \mathscr{A}, \{x_1, \cdots, x_k\} \subset \mathscr{X}, \{D_1, \cdots, D_k\} \subset \mathscr{D}, k \geqslant 1$.

**附注 6.3** MCSTRE 和 MCSTRE(+) 是两个不同的概念. $(\overrightarrow{X}, \overrightarrow{\xi})$ 是 MCSTRE, 但 $\{\overrightarrow{X}_0^\infty, \overrightarrow{\xi}_0^\infty)$ 未必是 MCSTRE(+), $(\overrightarrow{X}_0^\infty, \overrightarrow{\xi}_0^\infty)$ 是 MCSTRE(+) 但 $(\overrightarrow{X}, \overrightarrow{\xi})$ 未必是 MCTSRE. 反例如下:

令 $\mathscr{X} = \Gamma = \Theta = \{-1, 1\}, \eta$ 和 $\tau$ 是两个相互独立的具有公共分布 $P(\eta = -1) = P(\eta = 1) = \dfrac{1}{2}$ 的随机变量.

(1) 令 $X_{2n} = \eta,\ X_{2n+1} = \tau\ (n \in \mathbf{N}),\ \xi_{n,x} = \xi_{-n,x} = X_{n+1}\ \ (\forall n \in \mathbf{N}, x \in \mathscr{X}), p(\gamma, x, A) = \mathbf{1}_A(x)\ (\forall \gamma \in \Gamma, x \in \mathscr{X}, A \in \mathscr{A})$, 则 $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 是一个具有随机 Markov 核 $p(\gamma, x, A)$ 的 MCSTRE, 但 $\{\overrightarrow{X}_0^\infty, \overrightarrow{\xi}_0^\infty\}$ 不是一个具有随机 Markov 核的 MCSTRE(+).

(2) 令 $X_0 = \eta, X_n = \tau\ (\forall n \geqslant 1), \xi_{m,x} = \tau, (\forall x \in \mathscr{X}, m > 0), \xi_{0,x} \equiv 1\ (\forall x \in \mathscr{X}), \xi_{n,x} = \eta + \tau, (\forall x \in \mathscr{X}, n < 0), p(\gamma, x, A) = \mathbf{1}_A(1)$, 则 $(\overrightarrow{X}_0^\infty, \overrightarrow{\xi}_0^\infty)$ 是 MCSTRE(+), 但 $(\overrightarrow{X}, \overrightarrow{\xi})$ 不是一个 MCSTRE.

下面我们要给出 MCSTRE 的构造.

**定理 6.1** 给定 $(\mathscr{X}, \mathscr{A}), (\Gamma, \mathscr{D}), (\Theta, \mathscr{B})$ 如前, $\overrightarrow{\mathscr{A}} = \mathscr{A}^{\mathbf{N}}, \overrightarrow{\Theta} = \Theta^{\mathbf{Z}}, \overrightarrow{\mathscr{B}} = \mathscr{B}^{\mathbf{Z}}, \Phi$ 是可测空间 $(\mathscr{X} \times \overrightarrow{\Theta}, \mathscr{A} \times \overrightarrow{\mathscr{B}})$ 上的一个概率测度. 用 $x, \overrightarrow{x} = (x_n, n \in \mathbf{N})$ 和 $\overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z})$ 分别表示 $\mathscr{X}, \overrightarrow{\mathscr{X}}$ 和 $\overrightarrow{\Theta}$ 中的元素. $p(\gamma; x, A) \in \mathrm{RMK} - \Gamma$, 则存在一个概率空间 $(\Omega, \mathscr{F}, P)$ 及其上的 MCSTRE $\{\overrightarrow{X}, \overrightarrow{\xi}\}$ 具有随机 Markov 核 $p(\gamma; x, A)$, 使

$$(\overrightarrow{X}, \overrightarrow{\xi}): \Omega \longmapsto \overrightarrow{\mathscr{X}} \times \overrightarrow{\Theta},$$

而且满足

$$P((\overrightarrow{X}, \overrightarrow{\xi}) \in C) = \int_{\mathscr{X} \times \overrightarrow{\Theta}} \Phi(dx, d\overrightarrow{\theta}) \int_{\overrightarrow{\mathscr{X}}} P^x_{\overrightarrow{\theta}}(d\overrightarrow{x}) \mathbf{1}_C(\overrightarrow{x}, \overrightarrow{\theta}), \tag{6.3}$$

其中 $C \in \overrightarrow{\mathscr{A}} \times \overrightarrow{\mathscr{B}}$, $P^x_{\overrightarrow{\theta}}$ 是下式定义的 $\overrightarrow{\mathscr{A}}$ 上的概率测度:

$$\begin{aligned} &P^x_{\overrightarrow{\theta};}(\{\overrightarrow{x} \in \overrightarrow{\mathscr{X}} : x_0 \in A_0, x_1 \in A_1, \cdots, x_n \in A_n\}) \\ =& \int_{A_0} \delta(x, dx_0) \int_{A_1} p(\theta_{0,x_0}; x_0, dx_1) \cdots \int_{A_n} p(\theta_{n-1,x_{n-1}}; x_{n-1}, dx_n), \end{aligned} \tag{6.4}$$

其中 $\theta_i = (\theta_{i,x}, x \in \mathscr{X}) \in \Theta, \overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}, x \in \mathscr{X}, x_i \in \mathscr{X}, A_i \in \mathscr{A}, \theta_{i,x} \in \Gamma\ (0 \leqslant i \leqslant n), P^x_{\overrightarrow{\theta}}(A)$ 对固定的 $A \in \overrightarrow{\mathscr{A}}$ 是 $(x, \overrightarrow{\theta})$ 的关于 $\sigma$ 代数 $\mathscr{A} \times \overrightarrow{\mathscr{B}}$ 的可测函数. 注意: 由于 $p(\gamma; x, A)$ 对任何固定的 $A \in \mathscr{A}$ 是 $(\gamma, x)$ 的 $\mathscr{D} \times \mathscr{A}$ 可测函数, $\theta_{n,x} = \theta_n(x)$ 当 $n \in \mathbf{Z}$ 固定时是 $x$ 的 $\mathscr{A}/\mathscr{D}$ 可测函数, 所以 $p(\theta_{n,x}; x, A)$ 当 $n$ 和 $A$ 固定时是 $x$ 的 $\mathscr{A}$ 可测函数, 因此 (6.4) 式右方的积分是有定义的.

此外还有

(1) $\overrightarrow{X}$ 的 $P$ 边缘分布为

$$P \circ (\overrightarrow{X})^{-1}(A) = \int_{\mathscr{X} \times \overrightarrow{\Theta}} \Phi(dx, d\overrightarrow{\theta}) P^x_{\overrightarrow{\theta}}(A) \quad (A \in \overrightarrow{\mathscr{A}}); \tag{6.5}$$

(2) $\overrightarrow{\xi}$ 的 $P$ 边缘分布为

$$P \circ (\overrightarrow{\xi})^{-1}(B) = \Phi(\mathscr{X} \times B) \stackrel{\text{def.}}{=\!=} \pi(B) \quad (B \in \overrightarrow{\mathscr{B}}). \tag{6.6}$$

此定理的证明仿定理 1.1, 故其证明略去. 有兴趣的读者可做为习题验证之.

**定义 6.3** 设 $\overrightarrow{\Theta}$ 定义如前, $T$ 是由 $\overrightarrow{\Theta}$ 到 $\overrightarrow{\Theta}$ 的推移算子, 即是对任何 $\overrightarrow{\theta} = (\theta_n, n \in \mathbf{Z}) \in \overrightarrow{\Theta}, (T\overrightarrow{\theta})_n = \theta_{n+1}, p(\gamma; x, A) \in \mathrm{RMK} - \Gamma$. 对任何 $\overrightarrow{\theta} \in \overrightarrow{\Theta}, x_0 \in \mathscr{X}, A \in \mathscr{A}$, 令

$$\begin{aligned} p^{(n)}(\overrightarrow{\theta}; x_0, A) =& \int_{\mathscr{X}} p(\theta_{0,x_0}; x_0, dx_1) \cdots \\ &\int_{\mathscr{X}} p(\theta_{n-2,x_{n-2}}; x_{n-2}, dx_{n-1}) p(\theta_{n-1,x_{n-1}}; x_{n-1}, A), \end{aligned} \tag{6.7}$$

称 $p^{(n)}(\overrightarrow{\theta}; x_0, A)$ 为 $n$ 步随机转移函数, 有时也记之为 $p_n(\overrightarrow{\theta}; x_0, A)$. 这与确定环境的 Markov 核产生 $n$ 步转移函数是类似的.

再定义

$$\widehat{P}(n, (x_0, \overrightarrow{\theta}), F) = \int_{\mathscr{X}} p^{(n)}(\overrightarrow{\theta}; x_0, dy) \mathbf{1}_{(F)_y}(T^n \overrightarrow{\theta}) \tag{6.8}$$

$(n \geqslant 1, x_0 \in \mathscr{X}, \vec{\theta} \in \vec{\Theta}, F \in \mathscr{A} \times \vec{\mathscr{B}}, (F)_y = \{\vec{\theta} : (y, \vec{\theta}) \in F\}$ 是 $F$ 在 $y$ 的截口集. 记 $\widehat{P}((x_0, \vec{\theta}), F) \stackrel{\text{def.}}{=} \widehat{P}(1, (x_0, \vec{\theta}), F)$. 显然 $\widehat{P}(n, (x_0, \vec{\theta}), F)$ 是经典的 $n$ 步转移函数.

**定理 6.2** 沿用前面的符号. $\{(X_n, T^n \vec{\xi}), n \in \mathbf{N}\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的以 $(\mathscr{X} \times \vec{\Theta}, \mathscr{A} \times \vec{\mathscr{B}})$ 为状态空间的、以 $\widehat{P}(n, (x, \vec{\theta}), F)$ 为转移函数、以 $\Phi$ 为初始分布的时齐的 Markov 链 (见定义 1.3). 称之为斜积 Markov 链 (SPMC).

证明甚易, 读者可作为习题验证之.

**定理 6.3** 给定 $(\mathscr{X}, \mathscr{A})$、$(\Gamma, \mathscr{D})$, $(\Theta, \mathscr{B})$ 和 $p(\gamma; x, A)$ 如定理 6.1, $\vec{\Theta} = \Theta^{\mathbf{N}}$, $\vec{\mathscr{B}} = \mathscr{B}^{\mathbf{N}}, \vec{\mathscr{X}} = \mathscr{X}^{\mathbf{N}}, \vec{\mathscr{A}} = \mathscr{A}^{\mathbf{N}}, \Phi$ 是 $\mathscr{A} \times \vec{\mathscr{B}}$ 上的概率测度, $\pi(B) = \Phi(\mathscr{X} \times B)$. 恒存在概率空间 $(\Omega, \mathscr{F}, P)$ 及其上的 MCSTRE (+) $\{\vec{X}_0^\infty; \vec{\xi}_0^\infty\}$, 其随机 Markov 核是 $p$ 使 (6.3)~(6.6) 成立, 当然现在 $\vec{\theta} = (\theta_n, n \in \mathbf{N}) \in \vec{\Theta} \stackrel{\text{def.}}{=} \Theta^{\mathbf{N}}, \vec{\xi} = (\xi_n, n \in \mathbf{N})$.

证明甚易, 读者可作为习题验证之.

下面我们研究随机环境中的 Markov 性 (M) 的等价描述.

沿袭前面的符号. 引进条件:

$$(\mathrm{M}_0): \quad P(X_{n+1} \in A | \vec{X}_0^n, \vec{\xi}_{-\infty}^{\infty}) = p(\xi_{n, X_n}; X_n, A); \tag{6.9}$$

$$(\mathrm{M}_1): \quad P(X_0 \in A | \vec{\xi}_{-\infty}^{\infty}) = P(X_0 = A | \vec{\xi}_{-\infty}^{0}) \tag{6.10}$$

$$\begin{aligned}(\mathrm{M}_2): \quad & P(X_{n+1} \in A, \vec{\xi}_{n+1}^{\infty} \in G | \vec{\xi}_0^n, \vec{\xi}_{-\infty}^n) \\ & = p(\xi_{n, X_n}; X_n, A) P\left(\vec{\xi}_{n+1}^{\infty} \in G | \vec{\xi}_{-\infty}^n\right),\end{aligned} \tag{6.11}$$

其中 $n \in \mathbf{N}, A \in \mathscr{A}, G \in \vec{\mathscr{B}}_{n+1}^{\infty})$.

类似地, 在 (6.9)、(6.10)、(6.11) 中把 $\vec{\xi}_{-\infty}^{\infty}, \vec{\xi}_{-\infty}^{n}, \vec{\xi}_{-\infty}^{0}$ 代之以 $\vec{\xi}_0^{\infty}, \vec{\xi}_0^n, \xi_0$, 则相应的条件分别记为 $(M_0^+, M_1^+, M_2^+$.

**附注 6.4** 设 $\sigma$ 代数 $\mathscr{G}_1, \mathscr{G}_2 \subset \mathscr{F}, Y$ 是 $(\Omega, \mathscr{F}, P)$ 上的随机变量, $\mathscr{G}_1 \subset \mathscr{G}_2, \boldsymbol{E}(|Y|) < \infty, \boldsymbol{E}(Y|\mathscr{G}_2)\mathscr{G}_1$ 可测, 则 $\boldsymbol{E}(Y|\mathscr{G}_1) = \mathbf{E}(Y|\mathscr{G}_2)$.

**定理 6.4** 恒有: (M)⇔($\mathrm{M}_0$) 和 ($\mathrm{M}_1$)⇔($\mathrm{M}_2$).

在证明定理以前, 先证明:

**引理 6.1** 若 (M) 成立, 则下列等式成立:

$$(1)\quad P\left(\bigcap_{k=1}^{m}\{\xi_{n+k}\in B_k\}|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right)=P\left(\bigcap_{k=1}^{m}\{\xi_{n+k}\in B_k\}|\overrightarrow{\xi}_{-\infty}^{\,n}\right);\tag{6.12}$$

$$\begin{aligned}(2)\quad &P\left(X_{n+1}\in A,\bigcap_{k=1}^{m}\{\xi_{n+k}\in B_k\}|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right)\\ &=p(\xi_{n,X_n};X_n,A)P\left(\bigcap_{k=1}^{m}\{\xi_{n+k}\in B_k\}|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right),\end{aligned}\tag{6.13}$$

其中 $n\geqslant 0,m\geqslant 1,A\in\mathscr{A},B_k\in\mathscr{B}$.

**证** 在 (M) 中取 $A=\mathscr{X}$ 得知 $m=1$ 时 (6.12) 成立. 在 (M) 中应用 (6.12) (取 $m=1$), 发现 (6.13) 对 $m=1$ 也成立.

下面对 $m$ 作归纳法来证明 (6.12) 和 (6.13) 对一切正整数 $m$ 都成立.

设 (6.12) 和 (6.13) 对 $m$ 成立, 要证对 $m+1$ 也成立.

由 (6.12) 和 (6.13) 对 $m$ 成立知: 对 $(\overrightarrow{\Theta}_{-\infty}^n,\overrightarrow{\mathscr{B}}_{-\infty}^n)$ 上的任何有界可测函数 $g$, 有

$$\begin{aligned}&\boldsymbol{E}\left(\mathbf{1}_{\{X_{n+1}\in A\}}\cdot g(\overrightarrow{\xi}_{-\infty}^{\,n})\cdot\mathbf{1}_{\{\xi_{n+1}\in B_1,\cdots,\xi_{n+m}\in B_m\}}|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right)\\ &=p(\xi_{n,X_n};X_n,A)\boldsymbol{E}(g(\overrightarrow{\xi}_{-\infty}^{\,n})\mathbf{1}_{\{\xi_{n+1}\in B_1,\cdots,\xi_{n+m}\in B_m\}}|\overrightarrow{\xi}_{-\infty}^{\,n}).\end{aligned}$$

由单调类定理知对任何 $(\overrightarrow{\Theta}_{n+1}^{n+m},\overrightarrow{\mathscr{B}}_{n+1}^{n+m})$ 上的有界可测函数 $h(\theta_{n+1},\cdots,\theta_{n+m})$ 有

$$\begin{aligned}&\boldsymbol{E}\left(\mathbf{1}_{\{X_{n+1}\in A\}}\cdot g(\overrightarrow{\xi}_{-\infty}^{\,n})h(\xi_{n+1},\cdots,\xi_{n+m})|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right)\\ &=p(\xi_{n,X_n};X_n,A)\boldsymbol{E}(g(\overrightarrow{\xi}_{-\infty}^{\,n})h(\xi_{n+1},\cdots,\xi_{n+m})|\overrightarrow{\xi_{-\infty}^{n}}).\end{aligned}$$

再用单调类定理可知: 对 $(\overrightarrow{\Theta}_{-\infty}^{n+m},\overrightarrow{\mathscr{B}}_{-\infty}^{n+m})$ 上的任何有界可测函数 $f$, 有

$$\begin{aligned}&\boldsymbol{E}\left(\mathbf{1}_{\{X_{n+1}\in A\}}f(\overrightarrow{\xi}_{-\infty}^{\,n+m})|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{\,n}\right)\\ &=p(\xi_{n,X_n};X_n,A)\boldsymbol{E}(f(\overrightarrow{\xi}_{-\infty}^{\,n+m})|\overrightarrow{\xi}_{-\infty}^{\,n}).\end{aligned}\tag{6.14}$$

由 (6.12)、(6.13) 对 $m$ 成立及 (6.14) 得

$$\begin{aligned}
& P\left(X_{n+1}\in A, \bigcap_{k=1}^{m+1}\{\xi_{n+k}\in B_k\}|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n\right) \\
= & \boldsymbol{E}\left[\mathbf{1}_{\{X_{n+1}\in A\}}\mathbf{1}_{\{\xi_{n+1}\in B_1,\cdots,\xi_{n+m}\in B_m\}}\right. \\
& \left.\cdot P(\xi_{n+m+1}\in B_{m+1}|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^{n+m})|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n\right] \\
\overset{(6.12)}{=\!=\!=} & \boldsymbol{E}\left[\mathbf{1}_{\{X_{n+1}\in A\}}\mathbf{1}_{\{\xi_{n+1}\in B_1,\cdots,\xi_{n+m}\in B_m\}}\right. \\
& \left.\cdot P(\xi_{n+m+1}\in B_{m+1}|\overrightarrow{\xi}_{-\infty}^{n+m})|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n\right] \\
\overset{(6.14)}{=\!=\!=} & p(\xi_{n,X_n};X_n,A)\boldsymbol{E}\left[\mathbf{1}_{\{\xi_{n+1}\in B_1,\cdots,\xi_{n+m}\in B_m\}}\right. \\
& \left.\cdot P(\xi_{n+m+1}\in B_{m+1}|\overrightarrow{\xi}_{-\infty}^{n+m})|\overrightarrow{\xi}_{-\infty}^n\right] \\
= & p(\xi_{n,X_n};X_n,A)P\left(\bigcap_{k=1}^{m+1}\{\xi_{n+k}\in B_k\}|\overrightarrow{\xi}_{-\infty}^n\right).
\end{aligned} \tag{6.15}$$

在 (6.15) 中取 $A=\mathscr{X}$ 得知: 对 $m+1$, (6.12) 也成立. 由 (6.15) 及 (6.12) 对 $m+1$ 成立知 (6.13) 对 $m+1$ 也成立. 命题证毕.

下面我们用引理 6.1 来证明定理 6.4.

第一步. (M)⇒(M$_2$). 用引理 6.1 和单调类定理立即可证出 (M)⇒(M$_2$).

第二步. (M$_2$)⇒(M$_0$) 和 (M$_1$).

设 (M$_2$) 成立. 则 (M) 更成立, 从而 (6.13) 成立, 且

$$p(\xi_{n,X_n};X_n,A)=P(X_{n+1}\in A|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n) \tag{6.16}$$

由 (6.13)、(6.16) 并用单调类定理可知: $X_{n+1}$ 与 $\overrightarrow{\xi}_{n+1}^{\infty}$ 关于 $\sigma$ 代数 $\sigma(\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n)$ 条件独立. 故由第二章定理 5.2 有

$$P(X_{n+1}\in A|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^n)=P(X_{n+1}\in A|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^\infty). \tag{6.17}$$

由 (6.16)、(6.17) 得

$$p(\xi_{n,X_n};X_n,A)=P(X_{n+1}\in A|\overrightarrow{X}_0^n, \overrightarrow{\xi}_{-\infty}^\infty). \tag{6.18}$$

此即 (M$_0$) 成立.

在 (M$_2$) 中取 $A=\mathscr{X}, n=0$, 得

$$P(\overrightarrow{\xi}_1^\infty\in G|X_0, \overrightarrow{\xi}_{-\infty}^0)=P(\overrightarrow{\xi}_1^\infty\in G|\overrightarrow{\xi}_{-\infty}^0),$$

此即 $\overrightarrow{\xi}_1^\infty$ 与 $X_0$ 关于 $\sigma(\overrightarrow{\xi}_{-\infty}^0)$ 条件独立, 亦即 (M$_1$) 成立 (参见第二章定理 5.2).

第三步. (M$_0$) 和 (M$_1$)⇒(M).

易证: 由 $(\mathrm{M}_0)$ 及单调类定理可知, 对任何 $k<n$, 及 $\overrightarrow{\Theta}_{-\infty}^{n-1}\times\mathscr{X}$ 上的有界可测实值函数 $f$, 总有

$$\begin{aligned}&\boldsymbol{E}\left(f(\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},X_{k+1})|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n+1}\right)\\ =&\boldsymbol{E}(f(\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},X_{k+1}|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n}).\end{aligned}\tag{6.19}$$

事实上, 由 $(\mathrm{M}_0)$ 和附注 6.4 知: 任取 $A\in\mathscr{A},B\in\overrightarrow{\mathscr{B}}_{-\infty}^{n-1},k<n$, 有

$$\begin{aligned}&\boldsymbol{E}(\mathbf{1}_A(X_{k+1})\mathbf{1}_B(\overrightarrow{\xi}_{-\infty}^{n-1})|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n+1})\\ =&\ \mathbf{1}_B(\overrightarrow{\xi}_{-\infty}^{n-1})\boldsymbol{E}(\mathbf{1}_A(X_{k+1})|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n+1})\\ \xlongequal{(\mathrm{M}_0)}&\ \mathbf{1}_B(\overrightarrow{\xi}_{-\infty}^{n-1})p(\xi_{k,X_k};X_k,A)\\ =&\int_{\mathscr{X}}p(\xi_{k,X_k};X_k,dx)\mathbf{1}_A(x)\mathbf{1}_B(\overrightarrow{\xi}_{-\infty}^{n-1}).\end{aligned}\tag{6.20}$$

由 (6.20), 并由单调类定理, 再注意 $k<n$ 和附注 6.4, 得

$$\begin{aligned}&\boldsymbol{E}(f(\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},X_{k+1})|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n+1})\\ =&\int_{\mathscr{X}}p(\xi_{k,X_k};X_k,dx)f(\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},x)\\ =&\boldsymbol{E}(f(\cdots,\xi_{-1},\xi_0,\xi_1,\cdots,\xi_{n-1},X_{k+1})|\overrightarrow{X}_0^k,\overrightarrow{\xi}_{-\infty}^{n}).\end{aligned}$$

(6.19) 获证.

由 (6.19) (取 $k=n-1$) 有

$$\begin{aligned}&\boldsymbol{E}(\mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-1}\in A_{n-1},X_n\in A_n\}}|\overrightarrow{X}_0^{n-1},\overrightarrow{\xi}_{-\infty}^{n+1})\\ =&\ \mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-1}\in A_{n-1}\}}\boldsymbol{E}(\mathbf{1}_{\{X_n\in A_n\}}|\overrightarrow{X}_0^{n-1},\overrightarrow{\xi}_{-\infty}^{n+1})\\ \xlongequal{(6.19)}&\ \mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-1}\in A_{n-1}\}}\boldsymbol{E}(\mathbf{1}_{\{X_n\in A_n\}}|\overrightarrow{X}_0^{n-1},\overrightarrow{\xi}_{-\infty}^{n}).\end{aligned}\tag{6.21}$$

类似地, 由 $(\mathrm{M}_0)$、附注 6.4 和 (6.19) 有:

$$\begin{aligned}&\boldsymbol{E}\left[\boldsymbol{E}(\mathbf{1}_{\{X_0\in A_0,\cdots,X_n\in A_n\}}|\overrightarrow{X}_0^{n-1},\overrightarrow{\xi}_{-\infty}^{n+1})|\overrightarrow{X}_0^{n-2},\overrightarrow{\xi}_{-\infty}^{n+1}\right]\\ =&\ \boldsymbol{E}\left[\mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-1}\in A_{n-1}\}}\boldsymbol{E}(\mathbf{1}_{\{X_n\in A_n\}}|\overrightarrow{X}_0^{n-1},\overrightarrow{\xi}_{-\infty}^{n})|\overrightarrow{X}_0^{n-2},\overrightarrow{\xi}_0^{n+1}\right]\\ \xlongequal{(\mathrm{M}_0)}&\ \mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-2}\in A_{n-2}\}}\boldsymbol{E}\left[\mathbf{1}_{\{X_{n-1}\in A_{n-1}\}}\right.\\ &\left.\cdot p(\xi_{n-1,X_{n-1}};X_{n-1},A_n)|\overrightarrow{X}_0^{n-2},\overrightarrow{\xi}_{-\infty}^{n+1}\right]\\ \xlongequal{(6.19)}&\ \mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-2}\in A_{n-2}\}}\boldsymbol{E}\left[\mathbf{1}_{\{X_{n-1}\in A_{n-1}\}}\right.\\ &\left.\cdot p(\xi_{n-1,X_{n-1}};X_{n-1},A_n)|\overrightarrow{X}_0^{n-2},\overrightarrow{\xi}_{-\infty}^{n}\right]\\ \xlongequal{(\mathrm{M}_0)}&\ \mathbf{1}_{\{X_0\in A_0,\cdots,X_{n-2}\in A_{n-2}\}}\boldsymbol{E}\left[\mathbf{1}_{\{X_{n-1}\in A_{n-1},X_n\in A_n\}}|\overrightarrow{X}_0^{n-2},\overrightarrow{\xi}_{-\infty}^{n}\right].\end{aligned}\tag{6.22}$$

把上述叠代做 $n$ 次, 得到

$$
\begin{aligned}
&P\left(\bigcap_{k=0}^{n}\{X_k\in A_k\}|X_0,\overrightarrow{\xi}_{-\infty}^{n+1}\right)\\
=&P\left(\bigcap_{k=0}^{n}\{X_k\in A_k\}|X_0,\overrightarrow{\xi}_{-\infty}^{n}\right),\quad (n\geqslant 0,A_k\in\mathscr{A}).
\end{aligned}\tag{6.23}
$$

设 $g$ 和 $h$ 分别是 $(\overrightarrow{\Theta}_{-\infty}^{n},\overrightarrow{\mathscr{B}}_{-\infty}^{n})$ 和 $(\mathscr{X},\mathscr{A})$ 上的有界实值可测函数, 由 $(\mathrm{M}_1)$ 可证

$$
\begin{aligned}
&\boldsymbol{E}(g(\overrightarrow{\xi}_{-\infty}^{n})h(X_0)|\overrightarrow{\xi}_{-\infty}^{n+1})\\
=&\boldsymbol{E}(g(\overrightarrow{\xi}_{-\infty}^{n})h(X_0)|\overrightarrow{\xi}_{-\infty}^{n}).
\end{aligned}\tag{6.24}
$$

用 (6.24) 和单调类定理可证对 $(\overrightarrow{\Theta}_{-\infty}^{n}\times\mathscr{X},\overrightarrow{B}_{-\infty}^{n}\times\mathscr{A})$ 上的任何有界实值可测函数 $f$, 有

$$
\begin{aligned}
&\boldsymbol{E}(f(\overrightarrow{\xi}_{-\infty}^{n},X_0)|\overrightarrow{\xi}_{-\infty}^{n+1})\\
=&\boldsymbol{E}(f(\overrightarrow{\xi}_{-\infty}^{n},X_0)|\overrightarrow{\xi}_{-\infty}^{n}).
\end{aligned}\tag{6.25}
$$

视 (6.23) 右端为 (6.25) 中的 $f(\overrightarrow{\xi}_{-\infty}^{n},X_0)$, 并把 (6.23) 两边关于 $\sigma$ 代数 $\sigma(\overrightarrow{\xi}_{-\infty}^{n+1})$ 取条件期望, 并利用 (6.25) 式得

$$
P\left(\bigcap_{k=0}^{n}\{X_k\in A_k\}|\overrightarrow{\xi}_{-\infty}^{n+1}\right)=P\left(\bigcap_{k=0}^{n}\{X_k\in A_k\}|\overrightarrow{\xi}_{-\infty}^{n}\right).\tag{6.26}
$$

(6.26) 说明 $\{X_0,\cdots,X_n\}$ 与 $\xi_{n+1}$ 关于 $\sigma$ 代数 $\sigma(\overrightarrow{\xi}_{-\infty}^{n})$ 条件独立, 所以

$$
P\left(\xi_{n+1}\in B|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n}\right)=P\left(\xi_{n+1}\in B|\overrightarrow{\xi}_{-\infty}^{n}\right),\quad (\forall n\geqslant 0,B\in\mathscr{B}).\tag{6.27}
$$

由 (6.27) 和 $(\mathrm{M}_0)$ 和附注 6.4 得

$$
\begin{aligned}
&P(X_{n+1}\in A,\xi_{n+1}\in B|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n})\\
=\ &\boldsymbol{E}(\mathbf{1}_{\{\xi_{n+1}\in B\}}P(X_{n+1}\in A|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n+1})|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n})\\
=\ &\boldsymbol{E}(\mathbf{1}_{\{\xi_{n+1}\in B\}}P(\xi_{n,X_n},X_n,A)|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n})\\
=\ &p(\xi_{n,X_n};X_n,A)P(\xi_{n+1}\in B|\overrightarrow{X}_0^n,\overrightarrow{\xi}_{-\infty}^{n})\\
\overset{(6.27)}{=\!=}\ &p(\xi_{n,X_n};X_n,A)P(\xi_{n+1}\in B|\overrightarrow{\xi}_{-\infty}^{n}).
\end{aligned}\tag{6.28}
$$

$(\forall n\geqslant 0,A\in\mathscr{A},B\in\mathscr{B})$. 此即 (M) 成立. 定理证毕.

**定理 6.5** 恒有:

$(M^+)\Leftrightarrow(M_0^+)$ 和 $(M_1^+)\Leftrightarrow(M_2^+)$.

证明仿定理 6.4. 读者可作为习题验证之.

思考题: 请读者总结一下本章引进了哪些概率测度, 它们的作用是什么? 是为什么问题服务的? 对随机环境中的 Markov 链有几种观点去看待它? 它与经典的 (即环境是一个常数) Markov 链的本质区别在哪里?

## §7 习题及应用

1. 证明命题 3.5.
2. 证明命题 3.7.
3. 证明命题 4.4.
4. 证明命题 4.5.
5. 证明定理 5.2.
6. 证明定理 5.3.
7. 证明定理 6.2.
8. 证明定理 6.3.

9. 设 $V=(\mathscr{X},\mathscr{A};\Theta,\mathscr{B};p,\Phi)$ 是 $p-\Phi$ 链. 其中 $\mathscr{X}=\mathbf{Z},\mathscr{A}$ 是 $\mathscr{X}$ 中一切子集构成的 $\sigma$ 代数, $\Theta=[0,1],\mathscr{B}=\mathscr{B}[0,1]$ 是 [0,1] 上的 Borel $\sigma$ 代数, $\theta_i:\Theta^{\mathbf{Z}}\to\Theta$ 的坐标映射, $\pi(B)\overset{\text{def.}}{=\!=}\Phi(\mathscr{X}\times B),B\in\mathscr{B}^{\mathbf{Z}}$, 令 $\pi=\mathscr{L}^{\mathbf{Z}},\mathscr{L}$ 是 [0,1] 上的 Lebesgue 测度. 则 $\{\theta_i,i\in\mathbf{Z}\}$ 是 $(\Theta^{\mathbf{Z}},\mathscr{B}^{\mathbf{Z}},\mathscr{L}^{\mathbf{Z}})$ 上的独立同分布的、其公共分布为 [0,1] 上的均匀分布的随机变量序列. 令 $p(\theta;x,x+1)=\theta=1-p(\theta;x,x-1)$ $(\forall\theta\in\Theta,x\in\mathscr{X})$, 再令 $\boldsymbol{E}_\pi(\cdot)$ 是关于概率测度 $\pi$ 的期望算子. 设 $\{\overrightarrow{X},\overrightarrow{\xi}\}$ 是上述 $p-\Phi$ 链生成的 MCTRE. (称之为具有依时的独立同分布的环境中的 Bernoulli 随机徘徊.) 试证:

(i) $\boldsymbol{E}_\pi(\theta_i)=\dfrac{1}{2}$, $(\forall i\in\mathbf{Z}),\pi=\pi\circ T^{-1}$;

(ii) $\mathscr{X}$ 中的任一状态 $x$ 都是零常返状态;

(iii) $p(\theta;x,y)$ 满足 §4 中的 $0-1$ 律条件;

(iv) $\mathscr{X}$ 中任一状态 $x$ 的周期 $d_x$ 都是 2;

(v) $\mathscr{X}$ 可分解为:

$$\mathscr{X}=\mathscr{R}_0(1)=\mathscr{R}_0(1,1)\bigcup\mathscr{R}_0(1,2).$$

10. 设 $V=(\mathscr{X},\mathscr{A};\Theta,\mathscr{B};p,\Phi)$ 是 $p-\Phi$ 链, 其中 $\mathscr{X}=\{0,1,2,\cdots\},\mathscr{A}$ 是 $\mathscr{X}$ 中一切子集所成的 $\sigma$ 代数, $\Theta=[0,1],\mathscr{B}=\mathscr{B}([0,1]),p(\theta;x,0)=1-p(\theta;x,x+1)=q_x\theta$ $(\forall\theta\in\Theta,x\in\mathscr{X}),0<q_x<1.\mathscr{L},\{\theta_i,i\in\mathbf{Z}\},\pi,\boldsymbol{E}_\pi$ 如习题 9 中所定义.

$\{\vec{X},\vec{\xi}\}$ 是由上述 $p-\Phi$ 链生成的 MCTRE (称之为具有依时的独立同分布环境中的更新过程). 试证:

(i) $\boldsymbol{E}_\pi(\theta_i)=\dfrac{1}{2},\pi=\pi\circ T^{-1}$;

(ii) $0-1$ 律条件成立;

(iii) 对任何 $x\in\mathscr{X}$, 只要 $x$ 是常返状态, 则 $d_x=1$;

(iv) 若令

$$\delta=\sum_{i=0}^{\infty}\prod_{t=0}^{i}\left(1-\frac{1}{2}q_t\right),$$

$$\gamma=\prod_{i=0}^{\infty}\left(1-\frac{1}{2}q_i\right),$$

则

$$\mathscr{X}=\begin{cases}\mathscr{N}, & 当\ \delta=\infty,\gamma>0,\\ \mathscr{R}_0(1), & 当\ \delta=\infty,\gamma=0,\\ \mathscr{R}_+(1), & 当\ \delta<\infty.\end{cases}$$

11. 设 $V=(\mathscr{X},\mathscr{A};\Theta,\mathscr{B};p,\Phi)$ 是 $p-\Phi$ 链, 其中 $\mathscr{X},\mathscr{A},\Theta,\mathscr{B},\mathscr{L},\pi,\{\theta_i,i\in\mathbf{Z}\}$ 与习题 10 同, 而

$$p(\theta;0,x)=p(\theta;1,x)=k_x\theta\quad(x\geqslant 1,\theta\in\Theta),$$

$$p(\theta;0,0)=p(\theta;1,0)=1-\sum_{x=1}^{\infty}k_x\theta\quad(\theta\in\Theta),$$

$$p(\theta;n,x)=0\quad(\theta\in\Theta,n\geqslant 2,x<n-1),$$

$$p(\theta;n,x)=p(\theta;1,x-n+1)\quad(\theta\in\Theta,n\geqslant 2,x\geqslant n-1),$$

其中 $k_2+k_3+\cdots>0,k_x\geqslant 0,\sum\limits_{x=1}^{\infty}k_x<1,\sum\limits_{x=0}^{\infty}k_x=1$.$(\vec{X},\vec{\xi})$ 是 $p-\Phi$ 链产生的 MCTRE (称之为具有依时的独立同分布的环境的排队过程). 试证:

(i) $\boldsymbol{E}_\pi(\theta_i)=\dfrac{1}{2},\pi=\pi\circ T^{-1}$;

(ii) $0-1$ 律条件成立;

(iii) 对任何 $x\in\mathscr{X}$, 只要 $x$ 是常返状态, 则 $d_x=1$;

(iv) 若令 $K(\lambda)=\sum\limits_{x=0}^{\infty}\widetilde{k}_x\lambda^x$, 其中

$$\widetilde{k}_x=\frac{1}{2}k_x\quad(当\ x\geqslant 1),$$

$$\widetilde{k}_0=1-\frac{1}{2}\sum_{x=1}^{\infty}k_x,$$

则

$$\mathscr{X} = \begin{cases} \mathscr{N}, & \text{当 } \lim\limits_{\lambda\uparrow 1} \dfrac{dK(\lambda)}{d\lambda} > 1; \\ \mathscr{R}_0(1), & \text{当 } \lim\limits_{\lambda\uparrow 1} \dfrac{dK(\lambda)}{d\lambda} = 1; \\ \mathscr{R}_+(1), & \text{当 } \lim\limits_{\lambda\uparrow 1} \dfrac{dK(\lambda)}{d\lambda} < 1. \end{cases}$$

12. 设 $X$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的一个随机变量, $\boldsymbol{E}(|X|)<\infty,\Gamma\in\mathscr{F},\mathscr{G}$ 是 $\mathscr{F}$ 中的子 $\sigma$ 代数, 定义

$$\boldsymbol{E}(X|\Gamma,\mathscr{G}) = \begin{cases} \boldsymbol{E}(X\mathbf{1}_\Gamma|\mathscr{G})/\boldsymbol{E}(\mathbf{1}_\Gamma|\mathscr{G}), & \text{当分母非 } 0, \\ 0, & \text{反之.} \end{cases}$$

试讨论此类条件期望的性质 (与经典的条件期望 $\boldsymbol{E}(X|\mathscr{G})$ 类比).

此类条件期望在随机环境中的 Markov 链的研究中是有用的. 这时 $\mathscr{G}$ 通常是随机环境 $\overrightarrow{\xi}$ 所产生的 $\sigma$ 代数.

13. 令 $(\mathscr{X},\mathscr{A}),(\Gamma,\mathscr{D}),(\Theta,\mathscr{B}),(\overrightarrow{\Theta},\overrightarrow{\mathscr{B}})\Phi,\pi$ 及 $(\Omega,\mathscr{F},P)$ 如定理 6.1 中所定义, $(\overrightarrow{X},\overrightarrow{\xi})$ 也是定理 6.1 中所定义的具有 $p(\gamma)\in \mathrm{RMK}-\Gamma$ 的 MCSTRE. 设 $\mathscr{X}=\mathbf{Z}^d, d$ 为 $\geqslant 1$ 的整数, $\Gamma$ 是 $\mathscr{X}$ 上的全体离散的概率分布, $\mathscr{D}$ 是其中的弱收敛拓扑所构成的 $\sigma$ 代数. 对任何 $\gamma\in\Gamma,\gamma(y)$ 代表分布 $\gamma$ 在 $y$ 的质量. 如果下列条件成立:

$$p(\xi_{n,x};x,y)=\xi_{n,x}(y-x),\ \ P-\text{a.s.},$$

$(\forall n\in\mathbf{Z},x,y\in\mathscr{X})$, 则称 $(\overrightarrow{X},\overrightarrow{\xi})$ 是一个依时依空的随机环境中的随机徘徊 (RWSTRE).

(1) 试求出 $\overrightarrow{X}$ 的有限维联合分布;

(2) 算出 $p^{(n)}(\overrightarrow{\theta};x,y)\overset{\text{def.}}{=\!=}P(X_n=y|X_0=x,\overrightarrow{\xi}=\overrightarrow{\theta})$;

(3) 讨论当 $\{\xi_{n,x},n\in\mathbf{Z},x\in\mathscr{X}\}$ 是一族独立同分布的随机元且具有公共分布 $P\circ(\xi_{n,x})^{-1}=\mu$ 的特殊情形, 这时有何深入结果?

# 第九章 Brown 运动与多维正态分布

## §1 多维正态分布

在初等概率论中, 我们曾经学习过一维正态分布 (通常用 $N(\mu,\sigma^2)$ 表示, 其中 $\mu$ 是其期望, $\sigma^2$ 是方差), 它的密度函数是

$$p(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{\frac{-1}{2\sigma^2}(x-\mu)^2\right\},\quad x\in(-\infty,\infty)$$

它对应的特征函数为

$$f(t)=\exp\left\{it\mu-\frac{1}{2}\sigma^2t^2\right\},\quad t\in(-\infty,\infty).$$

在这一节中, 我们将简单地讨论一些多维正态分布. 暂时约定: 矩阵 (特别地向量) 右上方加一撇表示其转量.

**定义 1.1** 若 $d$ 维随机向量 $X\overset{\text{def.}}{=\!=}(X_1,\cdots,X_d)$ 具有密度函数

$$p(x_1,\cdots,x_d)=\frac{1}{(\sqrt{2\pi})^d|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(x-\mu)\Sigma^{-1}(x-\mu)\right),\qquad(1.1)$$

其中 $x=(x_1,\cdots,x_d),\mu=(\mu_1,\cdots,\mu_d)$ 是 $d$ 维实值行向量, $\Sigma$ 是 $d\times d$ 维对称正定方阵, $|\Sigma|$ 是 $\Sigma$ 之行列式, $\Sigma^{-1}$ 是 $\Sigma$ 之逆矩阵, 则称 $X$ 具有$d$ **维正态分布**, 简记为 $X\sim N_d(\mu,\Sigma)$.

**命题 1.1** (1.1) 式中定义之函数 $p(x_1,\cdots,x_d)$ 确实是密函数 (可测性是显然的), 即

$$\int_{-\infty}^{\infty}dx_1\cdots\int_{-\infty}^{\infty}dx_d(p(x_1,\cdots,x_d))=1.\qquad(1.2)$$

证明甚易, 读者可作为习题验证之.

恒记 $N(\mu,\Sigma)=N_1(\mu,\Sigma)$, 这时 $\mu$ 是常数, $\Sigma=\sigma^2$.

**命题 1.2** 若 $X=(X_1,\cdots,X_d)$ 具有 $d$ 维正态分布 $N_d(\mu,\Sigma)$, 则

(1) $X$ 的特征函数为

$$f(t_1,\cdots,t_d)=\exp\left\{it\mu'-\frac{1}{2}t\Sigma t'\right\}, \tag{1.3}$$

$$t=(t_1,\cdots,t_d),\quad \mu=(\mu_1,\cdots,\mu_d).$$

(2) $\boldsymbol{E}(X)=\mu, \boldsymbol{E}((X-\mu)(X-\mu)')=\Sigma,$ (1.4)

即是 $\mu$ 是 $X$ 的期望向量, $\Sigma$ 是 $X$ 的协方差矩阵.

**附注 1.1** 若 $\Sigma=(\sigma_{i,j},1\leqslant i,j\leqslant d),\sigma_{i,j}=\sigma_{j,i},\Sigma$ 正定 (从而 $|\Sigma|>0$). 若记 $\Sigma^{-1}=(\sigma_{i,j},1\leqslant i,j\leqslant d)$, 则由 $\Sigma$ 之对称性有:

$$b_{i,j}\overset{\text{def.}}{=\!=}\frac{1}{|\Sigma|}\Sigma_{j,i}=\frac{1}{|\Sigma|}\Sigma_{i,j}, \tag{1.5}$$

其中 $\Sigma_{i,j}$ 是行列式 $|\Sigma|$ 中对应于第 $i$ 行第 $j$ 列的代数余子式.

**例 1.1** 设 $(X_1,X_2)$ 服从二维正态分布 $N_2(\mu,\Sigma)$,

$$\Sigma=\begin{pmatrix}\sigma_{1,1} & \sigma_{1,2}\\ \sigma_{1,2} & \sigma_{2,2}\end{pmatrix}\quad \sigma_{i,j}=\boldsymbol{E}((X_i-\mu_i)(X_j-\mu_j))$$

$\rho_{1,2}=\sigma_{1,2}/\sqrt{\sigma_{1,1}\sigma_{2,2}}$, 则

$$\Sigma^{-1}=\frac{1}{\sigma_{1,1}\sigma_{2,2}-\sigma_{1,2}^2}\begin{pmatrix}\sigma_{2,2} & -\sigma_{1,2}\\ -\sigma_{1,2} & \sigma_{1,1}\end{pmatrix}\overset{\text{def.}}{=\!=}(b_{i,j},1\leqslant i,j\leqslant 2).$$

$(X_1,X_2)$ 的密度函数为

$$\begin{aligned}p(x_1,x_2)&=\frac{1}{2\pi\sqrt{\sigma_{1,1}\sigma_{2,2}-\sigma_{1,2}^2}}e^{-\frac{1}{2}\sum\limits_{i,j=1}^{2}(x_i-\mu_i)b_{i,j}(x_j-\mu_j)}\\ &=\frac{1}{2\pi\sqrt{\sigma_{1,1}\sigma_{2,2}(1-\rho_{1,2}^2)}}\cdot\exp\left\{-\frac{1}{2\left(1-\rho_{1,2}^2\right)}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{1,1}}}\right)^2\right.\right.\\ &\quad\left.\left.+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{2,2}}}\right)^2-2\rho_{1,2}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{1,1}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{2,2}}}\right)\right]\right\}.\end{aligned} \tag{1.6}$$

**命题 1.3** 若 $\Sigma$ 是 $d$ 阶正定方阵, 则 $\Sigma^{-1}$ 亦正定, 且对 $\Sigma$ 的任一特征值 $\lambda$ (对应的特征向量记为 $y_\lambda$), 必有 $\lambda>0,\dfrac{1}{\lambda}$ 为 $\Sigma^{-1}$ 的一个特征值, 对应的特征向量亦为 $y_\lambda$.

**证** $\Sigma$ 正定, $\lambda$ 为其一个特征值, 相应的特征向量为 $y_\lambda, y'_\lambda$ 为 $y_\lambda$ 之转置 (故 $y'_\lambda$ 是 $d$ 维行向量), 则由 $y_\lambda \neq 0$ 知

$$0 < y'_\lambda \Sigma y_\lambda = y'_\lambda \lambda y_\lambda = \lambda(y'_\lambda y_\lambda),$$

所以 $\lambda > 0$. 又由

$$y_\lambda = \Sigma^{-1}\Sigma y_\lambda = \lambda \Sigma^{-1} y_\lambda$$

得

$$\Sigma^{-1} y_\lambda = \frac{1}{\lambda} y_\lambda,$$

即 $\frac{1}{\lambda}$ 是 $\Sigma^{-1}$ 的一个特征值, 相应的特征向量亦为 $y_\lambda$.

由 $\Sigma$ 的正定性及 (1.5) 式知 $\Sigma^{-1}$ 为实对称矩阵. 而 $\Sigma^{-1}$ 的特征值 $\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}$ 均大于 0, 则由线性代数知, $\Sigma^{-1}$ 必正定, 命题 1.3 证毕.

**命题 1.4** *$d$ 维随机向量 $X \stackrel{\text{def.}}{=\!=} (X_1, \cdots, X_d)$ 服从 $d$ 维正态分布 $N_d(\mu, \Sigma)$ 的充分必要条件是: 对任何 $d$ 维常值行向量 $a, X$ 的线性组合 $aX'$ 服从一维正态分布 $N(a\mu', a\Sigma a')$.*

**证** 必要性. 设 $X$ 服从 $N_d(\mu, \Sigma)$, 则其特征函数为

$$f(t) = e^{it\mu' - \frac{1}{2}t\Sigma t'}, \quad t = (t_1, \cdots, t_d)$$

$\mu = (\mu_1, \cdots, \mu_d)$, 从而 $aX'$ 的特征函数为

$$\begin{aligned} \boldsymbol{E}(e^{iaX'\tau}) &= \boldsymbol{E}(e^{i(\tau a)X'}) \\ &= \exp[i\tau(a\mu') - \frac{1}{2}\tau^2(a\Sigma a')], \end{aligned}$$

此即 $aX'$ 服从 $N(a\mu', a\Sigma a')$.

充分性. 若对任何 $d$ 维常数行向量 $a, aX'$ 具有正态分布 $N(a\mu', a\Sigma a')$, 则 $aX'$ 的特征函数为

$$\varphi(\tau) = \exp\left[i\tau a\mu' - \frac{1}{2}\tau^2(a\Sigma a')\right], \tag{1.7}$$

亦即

$$\boldsymbol{E}(\exp[i(\tau a)X']) = \exp\left[i(\tau a)\mu' - \frac{1}{2}(\tau a)\Sigma(\tau a)'\right]. \tag{1.8}$$

在上式中取 $\tau = 1$. 由于 $a$ 是任意的常数行向量, 故可视之为特征函数中的自变量, 于是 (1.8) 变为:

$$\boldsymbol{E}(\exp[iaX']) = \exp\left[ia\mu' - \frac{1}{2}a\Sigma a'\right]. \tag{1.9}$$

(1.9) 说明 $X$ 服从 $N_d(\mu, \Sigma)$. 命题证毕.

**命题 1.5** 设 $X=(X_1,\cdots,X_d)$ 服从 $N_d(\mu,\Sigma)$, 则对任何 $r\times d$ 阶常数矩阵 $D$ 及任何 $r$ 维常数列向量 $a$, $DX'+a$ 均服从 $N(D\mu'+a,D\Sigma D')$.

**证** 设 $t=(t_1,\cdots,t_r)$, 则 $DX'+a$ 的特征函数为:

$$\begin{aligned}&\boldsymbol{E}(\exp(it(DX'+a)))\\&=e^{ita}\bullet e^{i(tD)\mu'-\frac{1}{2}(tD\Sigma(tD)')}\\&=e^{it(a+D\mu')}\bullet e^{-\frac{1}{2}(t(D\Sigma D')t')},\end{aligned}$$

此即 $DX'+a$ 服从下述 $r$ 维正态分布:

$$N_r(a+D\mu',D\Sigma D').$$

**命题 1.6** 设 $X=(X_1,\cdots,X_d)$ 服从 $N_d(\mu,\Sigma)$, 则 $X$ 的任一子向量 $(X_{i_1},\cdots,X_{i_r})$ (其中 $r\leqslant d,1\leqslant i_1<i_2<\cdots<i_r\leqslant d$) 服从正态分布 $N_r(\mu^{(r)},\Sigma^{(r)})$, 其中

$$\mu^{(r)}=\mu \text{ on } \{i_1,\cdots,i_r\},$$

$\Sigma^{(r)}=\Sigma$ on $\{i_1,\cdots,i_r\}\times\{i_1,\cdots,i_r\}$.

**证** 显然成立.

**命题 1.7** 设 $X=(X_1,\cdots,X_d)$ 服从 $N_d(\mu,\Sigma)$, 则 $\{X_1,\cdots,X_d\}$ 相互独立的充分必要条件是: $\Sigma=\mathrm{diag}(\sigma_{i,i},1\leqslant i\leqslant d)$ 是对角矩阵.

**证** 设 $X$ 服从 $N_d(\mu,\Sigma)$, 则 $X$ 的特征函数如 (1.3), 且由 (1.4) $X$ 的协方差矩阵为 $\Sigma$.

若 $\Sigma$ 是对角矩阵, 则由 (1.3) 知 $X$ 的特征函数为 $X_1,\cdots,X_d$ 各自的特征函数的乘积, 从而 $\{X_1,\cdots,X_d\}$ 相互独立.

若 $\{X_1,\cdots,X_d\}$ 相互独立, 则必有 $\boldsymbol{E}((X_i-\mu_i)(X_j-\mu_j))=0$ $(\forall i\neq j)$, 所以由 (1.4) 知: $\Sigma$ 是对角矩阵.

## §2 Brown 运动及其简单性质

本章常用 $N_d(\mu,\Sigma)$ 表示 $d$ 维正态分布, 其中 $\mu$ 是 $d$ 维实值向量, $\Sigma$ 是 $d\times d$ 维实值对称方阵. $\mu$ 是期望向量, $\Sigma$ 是协方差矩阵.

**定义 2.1** 设 $X=\{X_t:t\in\boldsymbol{T}=[0,b)\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的取值 $\mathbf{R}^d$ 上的随机过程, $b$ 可为 $\infty$. 若

(1) $X_t$ 服从正态分布 $N(0,tI)(\forall 0\leqslant t<b)$, 其中 $I$ 是 $d$ 维单位矩阵;

(2) $X$ 具有**独立增量**, 即对任何 $0 \leqslant t_0 < t_1 < \cdots < t_k < b, \{X_{t_i} - X_{t_{i-1}} : 1 \leqslant i < k\}$ 是相互独立的 $k$ 个增量;

(3) $\forall \omega \in \Omega, X(\bullet, \omega)$ 在 $[0, b)$ 上连续;

(4) $X_0 \equiv 0$,

则称 $X = \{X_t : t \in [0, b)\}$ 是**标准的始于 0 的 $d$ 维 Brown 运动**, 记之为 S.B$^d$.M.O..

对于取值于 $\mathbf{R}^d$ 中的随机变量 $X_t$, 有时记为 $X_t = (X_{t,1}, \cdots, X_{t,d})$.$|x|$ 表示 $\mathbf{R}^d$ 中的元 $x$ 的欧氏范数.

由命题 1.7, 若 $X_t$ 服从 $d$ 维正态分布 $N_d(\mu, \Sigma)$, 则 $X_{t,1}, \cdots, X_{t,d}$ 相互独立的充要条件是: $\Sigma$ 是对角矩阵.

**附注 2.1** S.B$^d$.M.O. 总是存在的. 这就是著名的 Lévy 定理. 限于篇幅, 本书不证明了. 有兴趣的读者请参看 [35] 第五章定理 2.1.

**命题 2.1** 设 $X = \{X_t : t \in [0, b)\}$ 是 S.B$^d$.M.O., 则 $\{X_{t,k} : t \in [0, b)\}$ 是 S.B$^1$.M.O., 而且 $X_{t,1}, \cdots, X_{t,d}$ 相互独立 $(\forall t \in [0, b))$. (记 $X_t = (X_{t,1}, \cdots, X_{t,d})$ 为 $d$ 维行向量.)

**证** 由于 $X_t$ 服从正态分布 $N(0, tI), tI$ 是对角矩阵, 所以 $X_{t,1}, \cdots, X_{t,d}$ 相互独立 $(\forall t \in [0, b))$. 显然 $X_{t,k}$ 服从正态分布 $N(0, t)$, 而 $\{X_{t,k} : t \in [0, b)\}$ 满足定义 2.1 中的条件 (2), (3), (4) 是显然的, 故 $\{X_{t,k} : t \in [0, b)\}$ 是 S.B$^1$.M.O..

**命题 2.2** 设 $X = \{X_t : t \in [0, b)\}$ 是 S.B$^d$.M.O., $L$ 是 $d \times d$ 阶实值标准正交矩阵, 即 $L' = L^{-1}, L$ 的行列式 $|L| = 1$ (此处 $L'$ 表 $L$ 之转置), 作变换

$$Y_t = (Y_{t,1}, \cdots, Y_{t,d}) = X_t L = (X_{t,1}, \cdots, X_{t,d}) L (t \in [0, b)),$$

则 $Y = \{Y_t : t \in [0, b)\}$ 也是 S.B$^d$.M.O..

**证** 由于 $X_t$ 服从 $d$ 维正态分布 $N_d(0, tI)$, 而 $Y_t = X_t L, L$ 是标准正交矩阵, 所以 $Y_t$ 亦服从 $d$ 维正态分布 $N_d(0, tI)$. 直接验证易知 $\{Y_t : t \in [0, b)\}$ 还满足定义 2.1 中的条件 (2), (3), (4), 命题得证.

**命题 2.3** 若 $\{X_t : t \in [0, b)\}$ 是 S.B$^d$.M.O., 则它具有平稳增量, 即 $X_{t+h} - X_{s+h}$ 与 $X_t - X_s$ 之分布一样为 $N_d(0, (t-s)I) (\forall 0 \leqslant s < t < b, 0 \leqslant s+h \leqslant t+h < b)$.

**证** 由于 $X_{t+h} = (X_{t+h} - X_{s+h}) + (X_{s+h} - X_0)$, 而 $X_{t+h} - X_{s+h}$ 与 $X_{s+h} - X_0$ 相互独立, 所以若令 $X_v = (X_{v,1}, \cdots, X_{v,d}), u = (u_1, \cdots, u_d) \in \mathbf{R}^d, u' = \begin{pmatrix} u_1 \\ \vdots \\ u_d \end{pmatrix}$

为 $u$ 之转置, 则有

$$\boldsymbol{E}(e^{iX_{t+h}\bullet u'}) = \boldsymbol{E}(e^{i(X_{t+h}-X_{s+h})u'})\boldsymbol{E}(e^{iX_{s+h}u'}),$$

从而由 $X_t$ 服从正态分布 $N_d(0,tI)$ 可知

$$\begin{aligned}\boldsymbol{E}(e^{i(X_{t+h}-X_{s+h})u'}) &= \boldsymbol{E}(e^{iX_{t+h}u'})/\boldsymbol{E}(e^{iX_{s+h}u'})\\ &= e^{-\frac{1}{2}u(t+h)u'}/e^{-\frac{1}{2}u(s+h)u'}\\ &= e^{-\frac{1}{2}u(t-s)u'},\end{aligned}$$

此即 $X_{t+h}-X_{s+h}$ 与 $X_t-X_s$ 的分布一样都是 $N_d(0,(t-s)I)$.

**命题 2.4** 设 $\{X_t : t\in[0,\infty)\}$ 是 S.B$^d$.M.O., 则对任何正数 $K$, 有

$$P\left(\bigcap_{t\in[0,\infty)}\{|X_t|\leqslant K\}\right)=0.$$

**证** 因为对任何 $\omega, X(\bullet,\omega)$ 在 $[0,\infty)$ 上连续, 所以若令 $D$ 为 $[0,\infty)$ 中全体有理数集, 则

$$\bigcap_{t\in[0,\infty)}\{|X_t|\leqslant K\}=\bigcap_{t\in D}\{|X_t|\leqslant K\}$$

是概率空间中的可测集. 又因为

$$\begin{aligned}P\left(\bigcap_{t\in[0,\infty)}\{|X_t|\leqslant K\}\right) &\leqslant P(|X_t|\leqslant K)\\ &= \int_{|x|\leqslant K}\left(e^{-\frac{|x|^2}{2t}}/(2\pi t)^{\frac{d}{2}}\right)dx\\ &\leqslant \int_{|x|\leqslant K}\frac{1}{(2\pi t)^{\frac{d}{2}}}dx \quad (t\in[0,\infty)),\end{aligned}$$

命 $t\to\infty$ 即得命题 2.4.

**命题 2.5** 设 $\{X_t : t\in[0,b)\}$ 是 S.B$^d$.M.O., 任取 $0=t_0<t_1<\cdots<t_k<b$, 则

$$(X_{t_1},X_{t_2},\cdots,X_{t_k})=(X_{t_1,1},\cdots,X_{t_1,d};X_{t_2,1},\cdots,X_{t_2,d};\cdots;X_{t_k,1},\cdots,X_{t_k,d})$$

服从 $dk$ 维正态分布.

**证** 由于 $X_{t_1}-X_{t_0},X_{t_2}-X_{t_1},\cdots,X_{t_k}-X_{t_{k-1}}$. 相互独立, 且 $X_{t_i}-X_{t_{i-1}}$ 服从 $d$ 维正态分布 $N_d(0,(t_i-t_{i-1})I)$, 所以 $X_{t_i,1}-X_{t_{i-1},1},\cdots,X_{t_i,d}-X_{t_{i-1},d}$ 相互独立且服从正态分布, 从而

$$\{X_{t_1,1}-X_{t_0,1},\cdots,X_{t_1,d}-X_{t_0,d};\cdots;X_{t_k,1}-X_{t_{k-1},1},\cdots,X_{t_k,d}-X_{t_{k-1},d}\}$$

相互独立, 且每个 $\{X_{t_i,l}-X_{t_{i-1},l}\}$ 皆服从正态分布, $i=1,\cdots,k,l=1,\cdots,d$. 但是若令

$$L_{d,k}=\underbrace{\begin{pmatrix} I & I & I & & I & I \\ 0 & I & I & \cdots & I & I \\ 0 & 0 & I & & I & I \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & & I & I \\ 0 & 0 & 0 & \cdots & 0 & I \end{pmatrix}}_{k\ \text{列}}\Bigg\}k\ \text{行}$$

其中 $I$ 是 $d$ 维单位矩阵, 则

$$\begin{aligned}&(X_{t_1},X_{t_2},X_{t_3},\cdots,X_{t_{k-1}},X_{t_k})\\ &=(X_{t_1}-X_{t_0},X_{t_2}-X_{t_1},X_{t_3}-X_{t_2},\cdots,X_{t_{k-1}}-X_{t_{k-2}},X_{t_k}-X_{t_{k-1}})\bullet L_{d,k},\end{aligned}$$

所以由命题 2.3 得知 $(X_{t_1},X_{t_2},X_{t_3},\cdots,X_{t_{k-1}},X_{t_k})$ 服从 $dk$ 维正态分布.

**命题 2.6**　*若* $\{X_t:t\in[0,b)\}$ *是* S.B$^1$.M.O., *则对任何* $s,t\in[0,b)$, *总有* $\mathrm{cov}(X_s,X_t)=s\wedge t.(s\wedge t\overset{\text{def.}}{=}\min(s,t).)$

**证**　由 $\{X_t:t\in[0,b)\}$ 具有独立增量, $X_t$ 服从正态分布 $N(0,t)$ 可知

$$\begin{aligned}\mathrm{cov}(X_s,X_t)&=\boldsymbol{E}(X_sX_t)\\ &=\mathrm{cov}(X_{s\wedge t},X_{s\vee t}-X_{s\wedge t})+\mathrm{cov}(X_{s\wedge t},X_{s\wedge t})\\ &=\mathrm{cov}(X_{s\wedge t},X_{s\wedge t})=\mathrm{var}(X_{s\wedge t})=s\wedge t.\end{aligned}$$

## §3　Brown 运动的轨道性质

在 §2 有关 S.B$^d$.M.O. 的定义中, 我们就给出了它的轨道的一个最基本的性质: 即一切轨道都是连续函数. 本节中, 我们将要讨论它的轨道的一些更深入的性质.

**定理 3.1**　*设* $\{X_t:t\in[0,\infty)\}$ *是概率空间* $(\Omega,\mathscr{F},P)$ *上的* S.B$^1$.M.O., *则存在* $A\in\mathscr{F},P(A)=0$, *当* $\omega\overline{\in}A$ *时* $X(\bullet,\omega)$ *在* $[0,\infty)$ *上无处可微.*

**证**　只需证明存在 $A\in\mathscr{F},P(A)=0$, 当 $\omega\overline{\in}A$ 时, $X(\bullet,\omega)$ 在 $[0,1)$ 无处可微即可.

任取定 $\omega\in\Omega$. 考虑 $X(\bullet,\omega):[0,1)\to\mathbf{R}$. 若存在 $s\in[0,1)$, 使 $X(\bullet,\omega)$ 在 $s$ 可微, 则必有

$$\lim_{t\downarrow s}\frac{X(t,\omega)-X(s,\omega)}{t-s}=X'(s,\omega)$$

为有限数, 所以

$$\begin{aligned}
&\{\omega\in\Omega\text{: 存在 } s\in[0,1)\text{, 使 } X(\bullet,\omega)\text{ 在 } s \text{ 可微 }\}\\
&\subset\bigcap_{k=1}^{\infty}\left\{\omega\in\Omega\text{: 存在 } s\in[0,1)\text{ 使 }\varlimsup_{t\downarrow s}\left|\frac{X(t,\omega)-X(s,\omega)}{t-s}\right|<k\right\}\\
&\subset\bigcup_{k=1}^{\infty}\{\omega\in\Omega\text{: 存在 } s\in[0,1),\delta=\delta(s,\omega)>0\text{, 使}\\
&\quad|X(t,\omega)-X(s,\omega)|\leqslant k|t-s|\text{, 当 } s\leqslant t<s+\delta\}.
\end{aligned}\tag{3.1}$$

但是, 当 $\frac{4}{n}<\delta, ns<i\leqslant ns+1$ 时, 有

$$\frac{i-1}{n}\leqslant s<\frac{i}{n}<\frac{i+1}{n}<\frac{i+2}{n}<\frac{i+3}{n}=\frac{4}{n}+\frac{i-1}{n}<\delta+s.\tag{3.2}$$

所以, 由

$$|X(t,\omega)-X(s,\omega)|\leqslant k|t-s|\quad(\text{当 } s\leqslant t<s+\delta)\tag{3.3}$$

及 (3.2) 式可推出

$$\begin{aligned}
\left|X\left(\frac{i+3}{n},\omega\right)-X\left(\frac{i+2}{n},\omega\right)\right|&\leqslant\left|X\left(\frac{i+3}{n},\omega\right)-X(s,\omega)\right|\\
&\quad+\left|X\left(\frac{i+2}{n},\omega\right)-X(s,\omega)\right|\\
&\leqslant k\left|\frac{i+3}{n}-s\right|+k\left|\frac{i+2}{n}-s\right|\\
&\leqslant k\bullet\frac{4}{n}+k\bullet\frac{3}{4}=\frac{7k}{n}.
\end{aligned}\tag{3.4}$$

仿之可证

$$\left|X\left(\frac{i+2}{n},\omega\right)-X\left(\frac{i+1}{n},\omega\right)\right|\leqslant\frac{5k}{n}<\frac{7k}{n},\tag{3.5}$$

$$\left|X\left(\frac{i+1}{n},\omega\right)-X\left(\frac{i}{n},\omega\right)\right|\leqslant\frac{3k}{n}<\frac{7k}{n}.\tag{3.6}$$

故当 $\frac{4}{n}<\delta, ns<i\leqslant ns+1$ 时,

$$\begin{aligned}
&|X(t,\omega)-X(s,\omega)|\leqslant k|t-s|\quad(s\leqslant t<s+\delta)\\
&\qquad\Rightarrow\left|X\left(\frac{j}{n},\omega\right)-X\left(\frac{j-1}{n},\omega\right)\right|\leqslant\frac{7k}{n}\quad(j=i+1,\cdots,i+3),
\end{aligned}$$

所以

$$\begin{aligned}
&\{\omega\in\Omega:|X(t,\omega)-X(s,\omega)|\leqslant k|t-s|,s\leqslant t<s+\delta\}\\
&\subset\bigcap_{n>\frac{4}{\delta}}\ \bigcup_{ns<i\leqslant ns+1}\ \bigcap_{i+1\leqslant j\leqslant i+3}\left\{\omega:\left|X\left(\frac{j}{n},\omega\right)-X\left(\frac{j}{n},\omega\right)\right|\leqslant\frac{7k}{n}\right\}.
\end{aligned}$$

因此

$$\begin{aligned}&\{\omega\in\Omega;\ \text{存在}\ s\in[0,1),\delta=\delta(s,\omega)>0,\ \text{使}\\&|X(t,\omega)-X(s,\omega)|\leqslant k|t-s|,\ \text{当}\ s\leqslant t<s+\delta\}\\\subset&\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\bigcup_{1\leqslant i\leqslant n+1}\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\omega:\left|X\left(\frac{j}{n},\omega\right)-X\left(\frac{j-1}{n},\omega\right)\right|\leqslant\frac{7k}{n}\right\},\end{aligned}\tag{3.7}$$

所以若令

$$A=\bigcup_{k=1}^{\infty}\bigcup_{m=1}^{\infty}\bigcap_{n=m}^{\infty}\bigcup_{1\leqslant i\leqslant n+1}\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\omega:\left|X\left(\frac{j}{n},\omega\right)-X\left(\frac{j-1}{n},\omega\right)\right|\leqslant\frac{7k}{n}\right\},$$

则有

$$\{\omega\in\Omega:\ \text{存在}\ s\in[0,1)\ \text{使}\ X(\cdot,\omega)\ \text{在}\ s\ \text{可微}\}\subset A.\tag{3.8}$$

但是

$$P(A)\leqslant\sum_{k=1}^{\infty}\sum_{m=1}^{\infty}P\left(\bigcap_{n=m}^{\infty}\bigcup_{1\leqslant i\leqslant n+1}\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\left|X\left(\frac{j}{n}\right)-X\left(\frac{j-1}{n}\right)\right|\leqslant\frac{7k}{n}\right\}\right),\tag{3.9}$$

由 $\{X(t):t\in[0,\infty)\}$ 具有独立平稳增量及 $X(t)\sim N(0,t)$ 可知:

$$\begin{aligned}&P\left(\bigcap_{n=m}^{\infty}\bigcup_{1\leqslant i\leqslant n+1}\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\left|X\left(\frac{j}{n}\right)-X\left(\frac{j-1}{n}\right)\right|\leqslant\frac{7k}{n}\right\}\right)\\\leqslant&\lim_{n\to\infty}\sup P\left(\bigcup_{1\leqslant i\leqslant n+1}\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\left|X\left(\frac{j}{n}\right)-X\left(\frac{j-1}{n}\right)\right|\leqslant\frac{7k}{n}\right\}\right)\\\leqslant&\lim_{n\to\infty}\sup\sum_{i=1}^{n+1}P\left(\bigcap_{i+1\leqslant j\leqslant i+3}\left\{\left|X\left(\frac{j}{n}\right)-X\left(\frac{j-1}{n}\right)\right|\leqslant\frac{7k}{n}\right\}\right)\\=&\lim_{n\to\infty}\sup\sum_{i=1}^{n+1}\prod_{j=i+1}^{i+3}P\left(\left|X\left(\frac{j}{n}\right)-X\left(\frac{j-1}{n}\right)\right|\leqslant\frac{7k}{n}\right)\\=&\lim_{n\to\infty}\sup\sum_{i=1}^{n+1}P\left(\left|X\left(\frac{1}{n}\right)\right|\leqslant\frac{7k}{n}\right)^3\\=&\lim_{n\to\infty}\sup(n+1)\left(\int_{|y|\leqslant\frac{7k}{n}}\frac{1}{\sqrt{2\pi\frac{1}{n}}}e^{-\frac{ny}{2}}dy\right)^3\end{aligned}$$

$$= \lim_{n\to\infty} \sup(n+1)\left(\int_{|y|<\frac{7k}{\sqrt{n}}} \sqrt{\frac{n}{2\pi}} e^{-\frac{y^2}{2}} \bullet \frac{1}{\sqrt{n}} dy\right)^3$$

$$\leqslant \lim_{n\to\infty} \sup(n+1)\left(\int_{|y|\leqslant\frac{7k}{\sqrt{n}}} dy\right)^3$$

$$= \lim_{n\to\infty} \sup(n+1)\left(\frac{14k}{\sqrt{n}}\right)^3 = 0. \tag{3.10}$$

由 (3.8), (3.9), (3.10) 式得定理 3.1.

**定理 3.2** 设 $\{X(t):t\in[0,\infty)\}$ 是 S.B¹.M.O., 令

$$Y_n = \sum_{k=1}^{2^n}\left[X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)\right]^2, \tag{3.11}$$

$$\lim_{n\to\infty} Y_n = 1, \quad L^2, \quad 且 \quad \text{a.s.}. \tag{3.12}$$

**证** 由命题 2.3 得知

$$\left[X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)\right] \sim N\left(0, \frac{1}{2^n}\right),$$

所以

$$\boldsymbol{E}(Y_n) = \sum_{k=1}^{2^n} \boldsymbol{E}\left(\left|X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)\right|^2\right) = 1.$$

因此

$$\begin{aligned}
&\boldsymbol{E}(|Y_n - 1|^2) = \operatorname{var}(Y_n)\\
&= \sum_{k=1}^{2^n} \operatorname{var}\left(\left(X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)\right)^2\right)\\
&= 2^{-2n}\sum_{k=1}^{2^n} \operatorname{var}\left(\left(\frac{X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)}{\sqrt{\frac{1}{2^n}}}\right)^2\right).
\end{aligned} \tag{3.13}$$

令

$$Z_{n,k} = \left(X\left(\frac{k}{2^n}\right) - X\left(\frac{k-1}{2^n}\right)\right) \Big/ \sqrt{\frac{1}{2^n}}, \tag{3.14}$$

$$Z_{n,k} \sim N(0,1), \quad \boldsymbol{E}(Z_{n,k}^2) = 1, \quad \boldsymbol{E}(Z_{n,k}^4) = 3,$$

所以

$$\operatorname{var}(Z_{n,k}^2) = \boldsymbol{E}(Z_{n,k}^4) - \boldsymbol{E}(Z_{n,k}^2)^2 = 3 - 1 = 2. \tag{3.15}$$

将 (3.14), (3.15) 式代入 (3.13) 式, 得

$$\boldsymbol{E}(|Y_n-1|^2)=2^{-2n}\bullet 2^n\bullet 2=2^{-n+1}.$$

因此 $\lim\limits_{n\to\infty}\boldsymbol{E}(|Y_n-1|^2)=0$, 即

$$\lim_{n\to\infty}Y_n=1,\quad L^2.$$

但是,

$$P(|Y_n-1|>\varepsilon)\leqslant\frac{\mathrm{var}(Y_n)}{\varepsilon^2}=\frac{2^{-n+1}}{\varepsilon^2},$$

所以由 Borel – Cantelli 引理得知 $\lim\limits_{n\to\infty}Y_n=1$, a.s..

**定理 3.3** 设 $\{X(t):t\in[0,\infty)\}$ 是 S.B$^1$.M.O.,$Y_n$ 如定理 3.2 所定义, 则

$$\sup_{n\geqslant 1}\sum_{k=1}^{2^n}\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|=\infty,\quad \text{a.s.}. \tag{3.16}$$

**证** 由 $Y_n$ 的定义知

$$Y_n\leqslant\left(\max_{1\leqslant k\leqslant 2^n}\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|\right)\sup_{n\geqslant 1}\sum_{k=1}^{2^n}\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|$$

但 $X(\bullet,\omega)$ 在 $[0,1]$ 上一致连续, 所以

$$\lim_{n\to\infty}\max_{1\leqslant k\leqslant 2^n}\left|X\left(\frac{k}{2^n}\right)-X\left(\frac{k-1}{2^n}\right)\right|\equiv 0.$$

因此, 由 $\lim\limits_{n\to\infty}Y_n=1$, a.s., 得知 (3.16) 式成立.

**命题 3.1** 设随机变量 $Y$ 服从正态分布 $N(0,\sigma^2),\boldsymbol{E}(|Y|^\alpha)=\rho\sigma^\alpha(\forall\alpha\in\mathbf{R})$, 其中

$$\rho=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|x|^\alpha e^{-\frac{x^2}{2}}dx.$$

**证** 直接计算立即可得命题 3.1.

**定义 3.1** 设 $\{X_t:t\in\boldsymbol{T}\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上取值于 $\mathbf{R}^d$ 的随机过程, $\boldsymbol{T}$ 可为开区间, 可为闭区间, 可为半开半闭区间, 可为有穷区间亦可为无穷区间. 若

(1) $X_t-X_s$ 服从正态分布 $N(0,\sigma^2(t-s)I)(s,t\in\boldsymbol{T},\sigma^2>0,I$ 是 $d$ 维单位矩阵, 0 是 $d$ 维向量);

(2) $X$ 具有独立增量, 即是对任何 $t_0<t_1<\cdots<t_k,t_i\in\boldsymbol{T}\{X_{t_i}-X_{t_{i-1}}:1\leqslant i\leqslant k\}$ 相互独立, 则称 $X=\{X_t:t\in\boldsymbol{T}\}$ 是 $d$ **维 Brown 运动**. 记之为 B$^d$.M..

若还有 $0 \in \boldsymbol{T}, X_0 \equiv 0$, 则称 $X$ 是始于 0 的 $d$ 维 Brown 运动, 记之为 $\mathrm{B}^d$.M.O..

**命题 3.2** 设 $X = \{X_t : t \in \boldsymbol{T}\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的 $\mathrm{B}^d$.M., 则存在 $X$ 的一个完全可分的修正 $\widetilde{X} = \{\widetilde{X}_t : t \in \boldsymbol{T}\}$, 它几乎一切轨道 $\widetilde{X}(\cdot, \omega)$ 都是连续的, 即存在 $A \in \mathscr{F}, P(A) = 1$, 使得对每个 $\omega \in A, \widetilde{X}(\cdot, \omega)$ 在 $\boldsymbol{T}$ 上都连续.

**证** 令 $X_t = (X_{t,1}, \cdots, X_{t,d})$, 对任何 $t \in \boldsymbol{T}$, 任何 $1 \leqslant k \leqslant d$, 都有

$$\lim_{s \to t} \boldsymbol{E}(|X_{t,k} - X_{s,k}|^2) = \lim_{s \to t} \sigma^2 |t - s| = 0,$$

故 $\{X_{t,k} : t \in [0, b)\}$ 是随机连续的, 从而由第六章定理 1.1 得知: 它存在一个完全可分的修正 $\{\widetilde{X}_{t,k} : t \in \boldsymbol{T}\}$. 又由命题 3.1 知存在常数 $\rho$ 使

$$\boldsymbol{E}(|\widetilde{X}_{t,k} - \widetilde{X}_{s,k}|^\alpha) = \rho \sigma^\alpha |t - s|^{\frac{\alpha}{2}} \quad (\forall \alpha > 0)$$

(因为 $(\widetilde{X}_{t,k} - \widetilde{X}_{s,k}) \sim N(0, \sigma^2 |t - s|)$), 所以用 [108] §3.2 定理 2 (取 $\alpha > 1$ 即可) 得知: 对几乎所有的 $\omega, \widetilde{X}_{t,k}(\omega)$ 是 $t$ 的连续函数. 命题 3.2 得证.

**附注 3.1** 对 $\mathrm{B}^d$.M.$X$ 而言, 相应的命题 2.1 至命题 2.6 都成立.

**附注 3.2** 类似于 S.$\mathrm{B}^d$.M.O.,$\mathrm{B}^d$.M. 当然恒存在.

**附注 3.3** 设 $X$ 是 $\mathrm{B}^1$.M., 类似的定理 3.1、定理 3.2、定理 3.3 都成立.

**附注 3.4** 由于命题 3.2 成立, 以后所言 $\mathrm{B}^d$.M., 都假定是它的完全可分的修正, 因而其几乎所有的轨道都是连续的. 为简单计. 不妨设其轨道都是连续的 (通过把概率空间 "净化", 总可做到这点).

**附注 3.5** 所谓标准的始于 0 的取值于 $\mathbf{R}^d$ 的 Brown 运动 S.$\mathrm{B}^d$.M.O., 就是一般的 Brown 运动 $\mathrm{B}^d$.M. 当 $X_0 \equiv 0, \sigma^2 = 1$ 的特殊场合.

前面我们研究 Brown 的轨道性质时, 仅研究了轨道的连续性、不可微性及无 "有界变差性" 等. 下面我们将较深入地研究一下 Brown 运动的轨道性质.

仍设 $\dim(\cdot), \varphi - m(\cdot)$ 是第一章定义 5.1 和定义 5.2 中所定义的 Hausdorff 维数与由 $\varphi$ 决定的 Hausdorff 测度.

**定义 3.2** 设 $\mu$ 是 $\mathbf{R}^d$ 上的具有紧支撑 $K$ 的有限的 Borel 测度 (当然具有可数可加性), $\alpha \geqslant 0$, 称

$$I_\alpha(\mu) \stackrel{\text{def.}}{=\!=} \int_{\mathbf{R}^d} \int_{\mathbf{R}^d} \frac{1}{|x - y|^\alpha} (\mu(dx) \mu(dy)) \tag{3.17}$$

为 $\mu$ 的 $\alpha-$ 能; 称

$$C_\alpha(K) \stackrel{\text{def.}}{=\!=} \sup \left\{ \frac{1}{I_\alpha(\mu)} : \mu \text{ 是具有紧支撑 } K^* \subset K \text{ 的 Borel 概率测度} \right\} \tag{3.18}$$

为 $K$ 的 $\alpha$–**容度**; 称

$$\dim_C(K) \stackrel{\text{def.}}{=\!=} \inf\{\alpha > 0 : C_\alpha(K) = 0\} = \sup\{\alpha > 0 : C_\alpha(K) > 0\}$$

为 $K$ 的**容度维数** (约定 $\frac{1}{\infty} = 0$).

众所周知: 由 Frostman 定理 (参见 [40] 第一章定理 4.5) 得

$$\dim_C(K) = \dim(K) \quad (\forall \text{ 紧集 } K \subset \mathbf{R}^d). \tag{3.19}$$

和以前一样, 仍用 $f(A)$ 表示函数 $f$ 在 $A$ 上之像集 ($A$ 是 $f$ 的定义域中任一子集), $f^{-1}(B)$ 表示函数 $f$ 在 $B$ 的逆像集 ($B$ 是 $f$ 的值域中的任一子集). 如果 $X = \{X(t) : t \in \boldsymbol{T}\}$ 是以 $(E, \mathscr{E})$ 为状态空间的随机过程, $A \subset \boldsymbol{T}, B \subset E$, 则定义 $X(A) \stackrel{\text{def.}}{=\!=} X(A)(\bullet), X(A)(\omega) \stackrel{\text{def.}}{=\!=} \{X(t,\omega) : t \in A\} (\forall \omega \in \Omega); X^{-1}(B) \stackrel{\text{def.}}{=\!=} X^{-1}(B)(\bullet), X^{-1}(B)(\omega) \stackrel{\text{def.}}{=\!=} \{t \in \boldsymbol{T} : X(t,\omega) \in B\}$. 称 $X(A), X^{-1}(B)$ 分别为 $X$ 在 $A$ 上之**像集**与在 $B$ 上之**逆像集**. 显然, $X(A)$ 和 $X^{-1}(B)$ 都是随机集. 有时记

$$X(A)(\omega) = X(A, \omega), \quad X^{-1}(B)(\omega) = X^{-1}(B, \omega).$$

**定理 3.4**　设 $X = \{X(t) : t \in [0,1]\}$ 是 $\mathrm{B}^d$.M.O., $d \geqslant 2$, 则 $P(\dim(X([0,1])) = 2) = 1$.

**证**　首先我们指出: 由 $X$ 的轨道的连续性可知:

对任何 $\omega, X([0,1])(\omega)$ 是 $\mathbf{R}^d$ 中的紧集, 所以, 由 Frostman 定理可知

$$\dim(X([0,1])) \equiv \dim_C(X([0,1])). \tag{3.20}$$

(1) 推证

$$\dim_C(X([0,1])) \geqslant 2, \quad \text{a.s.}. \tag{3.21}$$

为此, 只需证明对任何 $1 < \alpha < 2$, 有

$$C_\alpha(X([0,1])) > 0, \quad \text{a.s.},$$

亦即只需证明: 对几乎所有的 $\omega$, 存在 $X([0,1])$ 上的 Borel 概率测度 $\mu$, 使

$$I_\alpha(\mu) < \infty. \tag{3.22}$$

($C_\alpha$ 与 $I_\alpha$ 分别由 (3.18) 与 (3.17) 所定义.)

事实上, 由 $X_t - X_s$ 服从正态分布 $N(0, \sigma^2(t-s)I)$ 及命题 3.1 可知存在常数 $c$ 使

$$\boldsymbol{E}(|X_t - X_s|^{-\alpha}) = c|t-s|^{-\frac{\alpha}{2}}. \tag{3.23}$$

把 (3.23) 两边积分得

$$\int_{\Omega}\left(\int_0^1\int_0^1 |X_t - X_s|^{-\alpha} dsdt\right) dP = c\int_0^1\int_0^1 |t-s|^{-\frac{\alpha}{2}} dsdt < \infty.$$

所以

$$\int_0^1\int_0^1 |X_t - X_s|^{-\alpha} dsdt < \infty, \quad \text{a.s..} \tag{3.24}$$

固定任意 $\omega \in \Omega$, 令 $D(\omega) = X([0,1],\omega)$,

$$\mu_\omega(A) = \mathscr{L}_1(\{t \in [0,1] : X(t,\omega) \in A\}), \quad A \in \mathscr{B}(D(\omega)),$$

其中 $\mathscr{L}_1$ 是一维 Lebesgue 测度, $\omega \in \Omega$.

由于 $X(\bullet,\omega)$ 连续, 所以轨道 $X(\bullet,\omega)$ 可视为概率空间 $([0,1],\mathscr{B}([0,1]),\mathscr{L}_1)$ 上的取值于 $\mathbf{R}^d$ 的随机变量. 而 $\mu_\omega$ 就是 $X(\bullet,\omega)$ 的分布, 即

$$\mu_\omega = \mathscr{L}_1 \circ (X(\bullet,\omega))^{-1}. \tag{3.25}$$

($\mu_\omega$ 是 $X(\bullet,\omega)$ 的 "逗留时测度") 所以 $\mu_\omega$ 是 $D(\omega)$ 上的 Borel 概率测度 ($\forall \omega \in \Omega$).

由 (3.25) 和 (3.24) 可得

$$\begin{aligned} I_\alpha(\mu_\omega) &= \int_{D(\omega)}\int_{D(\omega)} |x-y|^{-\alpha}\mu_\omega(dx)\mu_\omega(dy) \\ &\bullet \int_{D(\omega)}\int_{D(\omega)} |x-y|^{-\alpha}\mathscr{L}_1(X(t,\omega)\in dx)\mathscr{L}_1(X(s,\omega)\in dy) \\ &= \int_0^1\int_0^1 |X(t,\omega)-X(s,\omega)|^{-\alpha} dsdt \\ &< \infty \ (\text{对几乎所有的 } \omega\in\Omega \text{ 及任意 } 1<\alpha<2). \end{aligned} \tag{3.26}$$

由 (3.26) 知

$$\dim_c(X([0,1])) \geqslant 2, \quad \text{a.s..} \tag{3.27}$$

(2) 推证

$$\dim_c(X([0,1])) \leqslant 2, \quad \text{a.s..} \tag{3.28}$$

为此, 只需证明对任意 $\alpha > 2$, 有

$$s^\alpha - m(X([0,1])) < \infty, \tag{3.29}$$

事实上, 任取 $\alpha > 2$, 有 $0 < \dfrac{1}{\alpha} < \dfrac{1}{2}$, 由 $\mathrm{B}^d$.M.O. 的 Hölder 连续性 (参见 [40] 第二章命题 1.3) 知: 对几乎所有的 $\omega$, 存在常数 $c > 0, \delta > 0$ (均不依赖 $\omega$), 使得

$$|X(t+h,\omega) - X(t,\omega)| \leqslant ch^{\frac{1}{\alpha}}$$
$$(\text{当 } 0\leqslant t\leqslant 1, t+h\leqslant 1, 0\leqslant h\leqslant\delta).$$

因此由上式即得

$$X([t,t+h]) \subset B(X(t), ch^{\frac{1}{\alpha}}),$$

其中 $B(x,r)$ 表示以 $x$ 为中心、以 $r$ 为半径之开球. 所以, 若取 $\frac{1}{m} < \delta$, 由上式立得

$$\begin{aligned}X([0,1]) &= \bigcup_{j=1}^{m} X\left(\left[\frac{j-1}{m}, \frac{j}{m}\right]\right) \\ &\subset \bigcup_{j=1}^{m} B\left(X\left(\frac{j-1}{m}\right), cm^{-\frac{1}{\alpha}}\right), \quad \text{a.s..}\end{aligned} \tag{3.30}$$

若令 $\mathscr{B}(d)$ 为 $\mathbf{R}^d$ 中一切开球, 则有常数 $K>0$ 使

$$\begin{aligned}s^{\alpha}-m(X([0,1])) &\leqslant K \lim_{m\to\infty} \inf\left\{\sum_{i=1}^{\infty}[\mathrm{diam}(B_i)]^{\alpha} : B_i \in \mathscr{B}(d),\right.\\ &\qquad \left.\bigcup_i B_i \supset X([0,1]), \mathrm{diam}(B_i) \leqslant 2cm^{-\alpha}\right\}\\ &\leqslant K \lim_{m\to\infty} \sum_{j=1}^{m}\left(\mathrm{diam}\left(B\left(X\left(\frac{j-1}{m}\right), cm^{-\frac{1}{\alpha}}\right)\right)\right)^{\alpha}\\ &= K \lim_{m\to\infty} m(2cm^{-\frac{1}{\alpha}})^{\alpha} = K(2c)^{\alpha} < \infty, \quad \text{a.s..}\end{aligned}$$

此即 (3.29) 成立, 从而 (3.28) 成立.

由 (3.20), (3.21), (3.28) 即得定理 3.4.

事实上, 我们还有比定理 3.4 更强的结果.

**定理 3.5**　*在定理 3.4 的条件下, 总有*

$$P(\dim(X(A)) = 2\dim A, \text{ 对 } [0,1] \text{ 中任何 Borel 集 } A \text{ 均成立}) = 1. \tag{3.31}$$

证明参见 [40] 第二章定理 2.3.

我们知道: 若 $X=\{X(t):t\in[0,\infty)\}$ 是 $\mathrm{B}^1$.M.O., 则对几乎所有的轨道 $X(\cdot,\omega)$ 来说, 它穿越任一水平线都有无穷多次, 即 $\{t\in[0,\infty):X(t,\omega)=x\}$ 是无穷集. 这个无穷集究竟有多 “大”? 下面的定理精确地回答了这一问题. 实际上, 它的结论所回答的问题比我们所提的还要广和深.

**定理 3.6**　*设 $X=\{X(t):t\in[0,\infty)\}$ 是 $\mathrm{B}^1$.M.O.,$B$ 是 $\mathbf{R}$ 中任一 Borel 集, 则*

$$P\left(\dim(X^{-1}(B)) = \frac{1+\dim(B)}{2}\right) = 1. \tag{3.32}$$

特别地, 当 $B=\{x\}$ 是 $\mathbf{R}$ 中的单点集时,$X^{-1}(B)=\{t:X(t)=x\}$ 就是 $X$ 穿越水平线 $x$ 的时间集合. (3.32) 告诉我们: 这个集合的 Hausdorff 维数是 $\frac{1}{2}$ (对几乎所有的轨道 $X(\bullet,\omega)$).

证明参见 [40] 第二章定理 4.3.

关于 Brown 运动轨道性质, 在此不再详细讨论了. 有兴趣的读者可参阅 [40] 及其后所列的参考文献.

## *§4 Wiener 空间及不变原理

本节讲述不变原理. 限于篇幅与本书的性质, 只讲述不变原理的本质、意义、结论及其应用, 而不给证明. 证明所需篇幅太大, 有兴趣从事这方面研究的读者, 可参看有关文献.

为讲述不变原理, 首先介绍 Wiener 空间.

**定义 4.1** 令 $\Omega$ 为全部定义在 $[0,\infty)$ 上取值于 $\mathbf{R}^d$ 的在零点函数值为 $0$($\mathbf{R}^d$ 中之零向量亦用 $0$ 表示) 的连续函数, $\Omega$ 中的元素用 $\omega$ 表示, 即

$$\Omega=\{\omega:\omega(\bullet):[0,\infty)\mapsto\mathbf{R}^d,\omega(\bullet)\text{ 连续},\omega(0)=0\}.$$

再令 $X_t$ 为 $\Omega$ 上的**坐标函数**, 即 $X_t(\omega)=\omega(t)(t\in[0,\infty),\omega\in\Omega)$, 有时亦记 $X_t$ 为 $X(t),X_t(\omega)$ 为 $X(t,\omega)$.

再令 $\mathscr{F}$ 是使一切 $X_t$ 可测的最小 $\sigma$ 代数, 即

$$\mathscr{F}=\bigvee_{t\in[0,\infty)}X_t^{-1}(\mathscr{B}^d)=\sigma(X^{-1}(B):t\in[0,\infty),B\in\mathscr{B}^d),$$

其中 $\mathscr{B}^d$ 是 $d$ 维欧氏空间 $\mathbf{R}^d$ 中的全体 Borel 集.

设 $\{\widehat{X}(t):t\in[0,\infty)\}$ 是概率空间 $(\widehat{\Omega},\widehat{\mathscr{F}},\widehat{P})$ 上的 $\mathrm{B}^d$.M.O., 对任何正整数 $m$, 任何 $B_i\in\mathscr{B}^d,i=1,\cdots,m$ 和 $0=t_0<t_1<t_2<\cdots<t_m$, 令 $J_m=\{t_1,\cdots,t_m\}$, 定义

$$\begin{aligned}&P_{J_m}([X(t_1)-X(t_0)]\in B_1,\cdots,[X(t_m)-X(t_{m-1})]\in B_m)\\=&\widehat{P}([\widehat{X}(t_1)-\widehat{X}(t_0)]\in B_1,\cdots,[\widehat{X}(t_m)-\widehat{X}(t_{m-1})]\in B_m)\\=&\prod_{i=1}^m\widehat{P}([\widehat{X}(t_i)-\widehat{X}(t_{i-1})]\in B_i)=\prod_{i=1}^m\widehat{P}(\widehat{X}(t_i-t_{i-1})\in B_i)\\=&\prod_{i=1}^m\int_{B_i}\left[\exp\left(-\frac{|x|^2}{2(t_i-t_{i-1})\sigma^2}\right)/(2\pi\sigma^2(t_i-t_{i-1}))^{\frac{d}{2}}\right]dx,\end{aligned}\tag{4.1}$$

其中, $x=(x_1,\cdots,x_d)\in\mathbf{R}^d, |x|^2=\sum\limits_{k=1}^{d}x_k^2, dx=dx_1,\cdots,dx_d$.

由 (4.1) 式所定义的集合函数 $P_{J_m}$ 显然可唯一地扩张到 $\sigma(X(t_1)-X(t_0), X(t_2)-X(t_1),\cdots,X(t_m)-X(t_{m-1}))=\sigma(X(t_1),\cdots,X(t_m))$ 上去, 而成为一概率测度 $P_{J_m}$.

令 $\boldsymbol{T}=[0,\infty),\varphi(\boldsymbol{T})$ 为 $\boldsymbol{T}$ 之一切有限子集, $E=\mathbf{R}^d,\mathscr{E}=\mathscr{B}^d,\mathscr{E}^S=\underset{t\in S}{X}\mathscr{E}_t,\mathscr{E}_t\equiv\mathscr{E},S\subset\boldsymbol{T}$. 任取 $J=\{t_1,\cdots,t_m\}\in\varphi(\boldsymbol{T}).A\in\mathscr{E}^J$, 定义

$$\widetilde{P}_J(A)=P_J(X_J^{-1}(A)),\tag{4.2}$$

其中 $X_J=(X(t_1),\cdots,X(t_m))$. 易证 $\{\widetilde{P}_J:J\in\varphi(\boldsymbol{T})\}$ 是 $(E,\mathscr{E})$ 上的一个投影测度系 (其定义参见 [35] 第四章定理 3.1). 再定义

$$\widetilde{P}(\pi_J^{-1}(A))=\widetilde{P}_J(A)\quad(A\in\mathscr{E}^J,J\in\varphi(\boldsymbol{T})),\tag{4.3}$$

则由 [35] 第四章定理 3.1, $\widetilde{P}$ 是全体柱集构成的代数

$$\mathscr{E}_0^{\boldsymbol{T}}=\{\Lambda:\Lambda=\pi_J^{-1}(A),A\in\mathscr{E}^J,J\in\varphi(\boldsymbol{T})\}$$

上的概率测度, 其中 $\pi_J$ 是由 [35] 第四章 (3.1) 式所定义的由 $E^{\boldsymbol{T}}$ 到 $E^J$ 的投影变换. 令

$$P(X_J^{-1}(A))=P_J(X_J^{-1}(A))=\widetilde{P}_J(A)=\widetilde{P}(\pi_J^{-1}(A))\tag{4.4}$$

$(\Lambda\in\mathscr{E}^J,J\in\varphi(\boldsymbol{T}))$. 由 $\widetilde{P}$ 是代数 $\mathscr{E}_0^J$ 上的概率测度易证 $P$ 是代数

$$\mathscr{F}_0\equiv\{M:M=X_J^{-1}(A),A\in\mathscr{E}^J,J\in\varphi(\boldsymbol{T})\}$$

上的概率测度, 故 $P$ 可唯一地扩张到 $\sigma$ 代数

$$\mathscr{F}=\bigvee_{t\in[0,\infty)}X_t^{-1}(\mathscr{B}^d)=\sigma(\mathscr{F}_0)$$

上去. 扩张后的概率测度仍用 $P$ 表示, 遂得概率空间 $(\Omega,\mathscr{F},P)$, 称之为 **Wiener 概率空间**, $P$ 称为 **Wiener 测度**.

Wiener 概率空间 $(\Omega,\mathscr{F},P)$ 上的坐标过程 $\{X(t):t\in[0,\infty),X(t,\omega)=\omega(t)\}$ 是一个 $\mathrm{B}^d$.M.O.. 因为由 (4.1) 式知定义 3.1 中的 (1) 和 (2) 成立, 由 $\Omega$ 的定义知 $X_0\equiv0$, 所以 $\{X(t):t\in[0,\infty)\}$ 是 $\mathrm{B}^d$.M.O..

**附注 4.1** 若取定义 4.1 中的 $\{\widehat{X}(t):t\in[0,\infty)\}$ 是概率空间 $(\widehat{\Omega},\widehat{\mathscr{F}},\widehat{P})$ 上的 S.$\mathrm{B}^d$.M.O., 则定义 4.1 中得到的 Wiener 空间 $(\Omega,\mathscr{F},P)$ 称为标准 Wiener 空间 ($P$ 为标准 Wiener 测度, 为了形象化本节用 $W$ 表示**标准 Wiener 测度**, 用 $(C[0,1],\mathscr{B}(C[0,1]),W)$ 表标准 Wiener 空间. 其上的坐标过程就是 S.$\mathrm{B}^d$.M.O.. 正因为如此, 有时称 $\mathrm{B}^d$.M. 为 Wiener 过程, 称 S.$\mathrm{B}^d$.M.O. 为**标准 Wiener 过程**.

现在我们简单地介绍不变原理.

在第五章中, 曾讲过中心极限定理 (简记为 CLT). 所谓一列随机变量满足中心极限定理, 意即此随机变量列的分布弱收敛到标准正态分布 $N(0,1)$. 最简单但也最常用的中心极限定理是: 若 $\{X_n, n\geqslant 1\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的一列二阶矩存在的独立同分布的随机变量, $\{S_n^*, n\geqslant 1\}$ 是 $\{X_n, n\geqslant 1\}$ 的正则化部分和, 即是

$$S_n=\sum_{k=1}^{n}X_k,\quad S_n^*=\frac{S_n-\boldsymbol{E}(S_n)}{\sqrt{\operatorname{var}(S_n)}}\quad(n\geqslant 1),$$

则随机变量列 $\{S_n^*, n\geqslant 1\}$ 满足 CL,T, 即

$$P\circ(S_n^*)^{-1}\overset{\mathrm{w}}{\Longrightarrow}N(0,1).$$

($\overset{\mathrm{w}}{\Longrightarrow}$ 表弱收敛, 定义见第三章).

那么什么叫不变原理呢? 设 $(C[0,1],\mathscr{B}(C[0,1]),W)$ 是前面定义的标准 Wiener 空间, $\{\{X^{(n)}(t),t\in[0,1]\},n\geqslant 1\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的一列实值随机过程, 简记随机过程 $\{X^{(n)}(t),t\in[0,1]\}$ 为 $X^{(n)}$. 显然 $X^{(n)}$ 可视为

$$\Omega\mapsto\mathbf{R}^{[0,1]}$$

的随机元.

若 $P\circ(X^{(n)})^{-1}\overset{\mathrm{w}}{\Longrightarrow}W$, (当 $n\to\infty$), 即对任何 $A\in\mathscr{B}(\mathbf{R}^{[0,1]})$, 只要 $A$ 为 $W$ 的连续集, 则有

$$\lim_{n\to\infty}P(X^{(n)}\in A)=W(A),$$

($A$ 为 $W$ 的连续集, 即 $W(A^\circ)=W(A^c)=W(A)$), 则称 $\{X^{(n)},n\geqslant 1\}$ 服从**不变原理**.

Donsker (参见 [20]) 在 1951 年证明了下述不变原理.

令

$\mathscr{X}=\{x: x(\cdot)$ 是定义在 $[0,1]$ 上的除了有限个点以外皆连续的实值函数 $\}$.

在 $\mathscr{X}$ 上定义距离:

$\rho(x,y)=\sup\limits_{t\in[0,1]}|x(t)-y(t)|, x,y\in\mathscr{X};$

$\mathscr{G}$ 是 $\rho$ 产生的距离拓扑;

$\mathscr{B}^{\mathscr{X}}\overset{\text{def.}}{=\!=}\mathscr{B}(\mathscr{G})$ 是 $\mathscr{G}$ 产生的 $\sigma$ 代数;

$C[0,1]$ 是定义在 $[0,1]$ 上的一切连续实值函数 $x$, 且 $x(0)=0$.

设 $\{u_i, i\geqslant 1\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的相互独立的具有公共分布的实值随机变量列, 而且 $\boldsymbol{E}(u_i)=0,\operatorname{var}(u_i)=1$. 令

$$S_n=\sum_{i=1}^{n}u_i\quad(n\geqslant 1).\tag{4.5}$$

对任意正整数 $n \geqslant 1$, 对任意 $n$ 个实数 $\tau_1, \cdots, \tau_n$, 定义 $\mathscr{X}$ 中的函数 $y_n$ 如下:

$$y_n(t; \tau_1, \cdots, \tau_n) = \begin{cases} \tau_1, & \text{当 } t = 0, \\ \tau_i, & \text{当 } t \in \left(\dfrac{i-1}{n}, \dfrac{i}{n}\right], i = 1, \cdots, n, \end{cases} \tag{4.6}$$

注意: 当 $n, \tau_1, \cdots, \tau_n$ 固定, $y_n(\bullet; \tau_1, \cdots, \tau_n) \in \mathscr{X}$.

设 $C_1 \subset C[0,1] \subset \mathscr{X}$ 且 $C_1$ 的 Wiener 测度 $W(C_1) = 1$.

设 $F: \mathscr{X} \mapsto \mathbf{R}$, 而且满足: $F$ 在 $C_1$ 上连续且 $\mathscr{B}^{\mathscr{X}}/\mathscr{B}(\mathbf{R})$ 可测.

令

$$X_{F,y}^{(n)}(t) = F\left(y_n\left(t; \frac{S_1}{\sqrt{n}}, \cdots, \frac{S_n}{\sqrt{n}}\right)\right) \quad (n \geqslant 1), \tag{4.7}$$

易见, 固定任一正整数 $n$ 及满足性质 (4.6) 的 $y_n$ 和上述 $F$,

$$X^{(n)} \overset{\text{def.}}{=\!=} \{X_{F,y}^{(n)}(t), t \in [0,1]\} \tag{4.8}$$

是一个随机过程.

**定理 4.1 (Donsker)** 设 $\{u_i, i \geqslant 1\}$ 是 $(\Omega, \mathscr{F}, P)$ 上的相互独立具有相同分布的随机变量列, 且 $\boldsymbol{E}(u_i) = 0, \mathrm{var}(u_i) = 1, \{X^{(n)}, n \geqslant 1\}$ 是由 (4.5)~(4.8) 所定义的随机过程列, 则 $\{X^{(n)}, n \geqslant 1\}$ 服从不变原理, 即

$$P \circ (X^{(n)})^{-1} \overset{\mathrm{w}}{\Longrightarrow} W. \tag{4.9}$$

Donsker 定理在不变原理发展的过程中, 是一个里程碑式的定理, 概括性强, 包含许多重要的有概率直观的极限定理作为它的特例. (参见文献 [20] 后面所列的参考文献.)

下述诸推论均沿袭定理 4.1 的符号.

**推论 4.1 (Kac – Erdös 定理)** 设 $\{u_i, i \geqslant 1\}$ 满足定理 4.1 中的条件, $S_n$ 是其部分和 $(n \geqslant 1)$, 则下述四个极限分布存在, 且不依赖 $u_i$ 之分布:

(1) $\xi_n^{(1)} = \max\{S_1, \cdots, S_n\}/\sqrt{n} \quad (n \geqslant 1)$;

(2) $\xi_n^{(2)} = \max\{|S_1|, \cdots, |S_n|\}/\sqrt{n} \quad (n \geqslant 1)$;

(3) $\xi_n^{(3)} = \dfrac{1}{n^2} \sum\limits_{i=1}^{n} S_i^2 \quad (n \geqslant 1)$;

(4) $\xi_n^{(4)} = \dfrac{1}{n^{3/2}} \sum\limits_{i=1}^{n} |S_i| \quad (n \geqslant 1)$.

**证** 在定理 4.1 中分别取 $F(x)$ 为

$$F^{(1)}(x) = \sup_{t\in[0,1]} x(t) \quad (x \in \mathscr{X});$$
$$F^{(2)}(x) = \sup_{t\in[0,1]} |x(t)| \quad (x \in \mathscr{X});$$
$$F^{(3)}(x) = \int_0^1 x(t)^2 dt \quad (x \in \mathscr{X});$$
$$F^{(4)}(x) = \int_0^1 |x(t)| dt \quad (x \in \mathscr{X}),$$

则可知

$$P \circ (\xi_n^{(k)})^{-1} = P \circ \left(F^{(k)}\left(y_n\left(t; \frac{S_1}{\sqrt{n}}, \cdots, \frac{S_n}{\sqrt{n}}\right)\right)\right)^{-1}$$
$$(k = 1, 2, 3, 4, n \geqslant 1).$$

所以由定理 4.1 立得推论 4.1 成立.

作者在文献 [41] 中把 Donsker 定理的固定加项的部分和 $S_n$ 换为随机加项的部分和 $S_{Z_n}$, 亦得到相应的不变原理, 其中 $\{Z_n, n \geqslant 1\}$ 是取正整数的随机变量列.

**定理 4.2 (见 [41])** 设 $\{u_i, i \geqslant 1\}$ 和 $\{Z_n, n \geqslant 1\}$ 是 $(\Omega, \mathscr{F}, P)$ 上的两个随机变量列, $\{u_i\}$ 相互独立且有相同的分布, $\{Z_n\}$ 取正整数, 而且

$$\lim_{n\to\infty} \frac{Z_n}{a_n} = Z, \quad P.$$

(其中 $\{a_n\}$ 是趋于 $\infty$ 的正实数列, $P(Z > 0) = 1, Z$ 与 $\{u_i\}$ 独立.) $F$ 满足定理 4.1 中的条件, $y_n$ 亦如 (4.6) 所定义. 令

$$X_{F,y}^{(Z_n)}(t) = F\left(y_{Z_n}\left(t; \frac{S_1}{\sqrt{Z_n}}, \cdots, \frac{S_{Z_n}}{\sqrt{Z_n}}\right)\right), \quad n \geqslant 1,$$
$$\widetilde{X}_{F,y}^{(Z_n)}(t) = F\left(y_{Z_n}\left(t; \frac{S_1}{\sqrt{a_n Z}}, \cdots, \frac{S_{Z_n}}{\sqrt{a_n Z}}\right)\right), \quad n \geqslant 1,$$
$$X^{(Z_n)} = \left\{X_{F,y}^{(Z_n)}(t), t \in [0,1]\right\}, \quad n \geqslant 1,$$
$$\widetilde{X}^{(Z_n)} = \left\{\widetilde{X}_{F,y}^{(Z_n)}(t), t \in [0,1]\right\}, \quad n \geqslant 1,$$

则

$$P \circ \left(X^{(Z_n)}\right)^{-1} \xrightarrow{\text{w}} W;$$
$$P \circ \left(\widetilde{X}^{(Z_n)}\right)^{-1} \xrightarrow{\text{w}} W.$$

下面我们简单地介绍一下 Markov 链的不变原理. 本段有关 Markov 链的基本概念均沿袭第六章.

设 $\{x_n, n = 0, 1, \cdots\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一个时齐的可数状态的 Markov 链, 其状态空间 $E$ 由一个正常返类所组成. $f$ 是定义在 $E$ 上的任一实值函数, 令

$$y_i = f(x_i), \quad S_n = \sum_{i=0}^{n} y_i.$$

再设 $\tau_n(i)$ 是 $\{x_u, u \geqslant 0\}$ 第 $n$ 次返回状态 $i$ 的时刻, 再令 $\rho_n(i) = \tau_{n+1}(i) - \tau_n(i)$, 而 $l(n, i)$ 是由关系

$$\tau_{l(n,i)}(i) \leqslant n < \tau_{l(n,i)+1}(i) \tag{4.10}$$

所确定的唯一的非负整值的随机变量. 取 $i$ 为任一固定的状态. 故可把 $i$ 略去不写, 以免符号过于复杂. 令

$$Y' = \sum_{j=0}^{\tau_1 - 1} y_j, \quad Y''(n) = \sum_{j=n+1}^{\tau_{l(n)+1}-1} y_j, \tag{4.11}$$

$$Y_\nu = \sum_{j=\tau_\nu}^{\tau_{\nu+1}-1} y_j, \tag{4.12}$$

则

$$S_n = Y' + \sum_{\nu=1}^{l(n)} Y_\nu + Y''(n). \tag{4.13}$$

再设 $\boldsymbol{E}(Y_\nu)$ 存在, 并令 $\mu = \boldsymbol{E}(Y_\nu)/\boldsymbol{E}(\rho_\nu), \overline{f} = f - \mu, z_j = \overline{f}(x_j) = y_j - \mu$. 再令

$$Z_\nu = \sum_{j=\tau_\nu}^{\tau_{\nu+1}-1} z_j, \quad U_\nu = \sum_{j=\tau_\nu}^{\tau_{\nu+1}-1} |z_j|, \tag{4.14}$$

则有

$$\begin{aligned} S_n - \mu n &= \sum_{j=0}^{\tau_1-1} z_j + \sum_{\nu=1}^{l(n)} Z_\nu - \sum_{j=n+1}^{\tau_{l(n)+1}-1} z_j \\ &\overset{\text{记作}}{=\!=} Z' + \sum_{\nu=1}^{l(n)} Z_\nu - Z''(n) \end{aligned} \tag{4.15}$$

再设 $\boldsymbol{E}(U_\nu^2) < \infty, 0 < \operatorname{var}(Z_\nu) = \boldsymbol{E}(Z_\nu^2) \overset{\text{def.}}{=\!=} \sigma^2 < \infty, \mathit{\Pi} = (\pi_{i,j}, i, j \in E)$ 是 $\{x_n, n = 0, 1, \cdots\}$ 的遍历极限. (参见第六章定理 5.3 中的 (5.6) 式). 再令

$$B = \sigma^2 \pi_{i,i}, \quad S_r^*(n) = \frac{S_r - r\mu}{\sqrt{nB}}, \quad W_\nu = \frac{Z_\nu}{\sigma}. \tag{4.16}$$

(注意: 由于 $\{x_n, n=0,1,\cdots\}$ 的状态空间 $E$ 由一个正常返类所构成, 所以由第六章定理 5.3 知 $\pi_{i,i}>0$, 从而 $B>0$, 故 (4.12) 式中诸量有定义.)

由 [13] P.78 定理 3 知 $\{W_\nu, \nu \geqslant 1\}$ 是相互独立的具有公共分布的随机变量序列, 而且 $\boldsymbol{E}(W_\nu)=0, \mathrm{var}(W_\nu)=1$.

对任何正整数 $n$ 和 $r$ 及任何正数 $b$, 令

$$Q_n=\sum_{\nu=1}^{n} W_\nu, \quad Q_r^*(b)=\frac{\sum\limits_{\nu=1}^{r} W_\nu}{\sqrt{b}}. \tag{4.17}$$

如 Donsker 定理中的符号, 仍令

$\mathscr{X}=\{\xi:\xi(\cdot)$ 是定义在 $[0,1]$ 上的、除了有限个跳跃点以外均连续的实值函数 $\}$;

$$\rho(\xi,\eta)=\sup_{t\in[0,1]}|\xi(t)-\eta(t)|, \quad \xi,\eta\in\mathscr{X};$$

$\mathscr{G}$ 是 $\rho$ 产生的距离拓扑;

$\mathscr{B}^{\mathscr{X}}=\mathscr{B}(\mathscr{G})$ 是 $\mathscr{G}$ 产生的 Borel $\sigma$ 代数;

$C([0,1])$ 是一切定义在 $[0,1]$ 上的实值连续函数 $x$ 且 $x[0]=0$.

$(C([0,1],\mathscr{B}(C[0,1]),W)$ 是前面定义的标准 Wiener 空间. $C_1\subset C([0,1])$, $W(C_1)=1$.

再令

$$I_{k,j}=\left(\frac{j-1}{k},\frac{j}{k}\right], \quad j=1,2,\cdots,k,$$

$$u_n(t;\lambda_1,\cdots,\lambda_n)=\begin{cases}\lambda_1, & t=0,\\ \lambda_j, & t\in I_{n,j}, j=1,\cdots,n,\end{cases} \tag{4.18}$$

显然, 对任意固定的正整数 $n$, 及实数 $\lambda_1,\cdots,\lambda_n, u_n(\cdot;\lambda_1,\cdots,\lambda_n)\in\mathscr{X}$.

**定理 4.3 ([42])** 在上面诸假设下, 我们有:

(1) 设 $F(\xi)$ 是定义在 $\mathscr{X}$ 上的有界泛函, 而且对拓扑 $\mathscr{G}$ 来说, 在 $C_1$ 上连续, 对 $\mathscr{B}^{\mathscr{X}}$ 来说是可测的, 则

$$\begin{aligned}&\lim_{n\to\infty}\int_\Omega F(u_n(t;S_1^*(n),\cdots,S_n^*(n)))P(d\omega)\\ &=\int_{C([0,1])}F(\xi)W(d\xi);\end{aligned} \tag{4.19}$$

(2) 设 $F(\xi)$ 是定义在 $\mathscr{X}$ 上的泛函 (不一定有界), 而且对拓扑 $\mathscr{G}$ 来说在 $C_1$

上连续, 对 $\mathscr{B}^{\mathscr{X}}$ 来说可测, 则

$$\begin{aligned}&\lim_{n\to\infty}\int_{\Omega}\exp(i\lambda F(u_n(t;S_1^*(n),\cdots,S_n^*(n))))P(d\omega)\\=&\int_{C([0,1])}\exp(i\lambda F(\xi))W(d\xi);\end{aligned}\tag{4.20}$$

(3) 设 $F(\xi)$ 满足 (2) 中的条件, 则对 $W(F(\xi)\leqslant\lambda)$ 的任何一个连续点 $\lambda$, 都有

$$\begin{aligned}&\lim_{n\to\infty}P(F(u_n(t;S_1^*(n),\cdots,S_n^*(n)))\leqslant\lambda)\\=&W(F(\xi)\leqslant\lambda),\end{aligned}\tag{4.21}$$

即下列随机过程列:

$$H_n\stackrel{\text{def.}}{=\!=}\{F(u_n(t;S_1^*(n),\cdots,S_n^*(n))),t\in[0,1]\},\quad n=1,2,\cdots$$

的分布 $P\circ H_n^{-1}$ 弱收敛到标准 Wiener 测度 $W$.

证明可参见 [42].

## §5 习题及应用

1. 设 $p$ 维随机向量 $X=\begin{pmatrix}X_1\\\vdots\\X_p\end{pmatrix}$ 服从 $p$ 维正态分布 $N(0,\Sigma)$, 即 $X$ 的期望向量 $\mu=\boldsymbol{E}(X)=0$, 协方差矩阵为 $\Sigma$. 记 $\Sigma=(\sigma_{i,j},i,j=1,2,\cdots,p)$. 试证明:

(1) 对任何不相等的 $i,j,k$, 有

$$\boldsymbol{E}(X_iX_jX_k)=0;$$

(2) 对任何不相等的 $i,j,k,l$, 有

$$\boldsymbol{E}(X_iX_jX_kX_l)=\sigma_{i,j}\sigma_{k,l}+\sigma_{i,k}\sigma_{j,l}+\sigma_{i,l}\sigma_{j,k}.$$

2. 设 $X=\begin{pmatrix}X_1\\X_2\end{pmatrix}$ 的联合密度函数为

$$f(x_1,x_2)=\begin{cases}\dfrac{1}{\pi}\cdot\exp\left(-\dfrac{x_1^2+x_2^2}{2}\right), & \text{当 } x_1x_2>0,\\0, & \text{反之},\end{cases}$$

试证: $X_1$ 和 $X_2$ 的边缘分布函数都是正态分布. (注意: 显然 $X$ 不服从二维正态分布.)

3. 设 $X_1,\cdots,X_n$ 为相互独立的随机变量, $X_i$ 服从正态分布 $N(\mu_i,\sigma_i^2)(i=1,\cdots,n)$, 再令

$$Y=\sum_{i=1}^{n}\sigma_i^{-2}X_i\Big/\sum_{i=1}^{n}\sigma_i^{-2},$$

$$Z=\sum_{i=1}^{n}\sigma_i^{-2}(X_i-Y)^2,$$

试证 $Y$ 与 $Z$ 相互独立.

4. 设 $X=\begin{pmatrix}X_1\\X_2\\X_3\end{pmatrix}$ 服从三元正态分布 $N_3(\mu,\Sigma)$, 其中

$$\mu=\begin{pmatrix}2\\-3\\1\end{pmatrix},\quad \Sigma=\begin{pmatrix}1&1&1\\1&3&2\\1&2&2\end{pmatrix},$$

试求 $3X_1-2X_2+X_1$ 的分布函数.

5. 设 $X_1$ 服从标准正态分布 $N(0,1)$, 令

$$X_2=\begin{cases}-X_1, & \text{当 } -1\leqslant X_1\leqslant 1,\\ X_1, & \text{反之},\end{cases}$$

试证:

(1) $X_2$ 也服从标准正态分布 $N(0,1)$;

(2) $(X_1,X_2)$ 不服从二元正态分布.

6. 请讨论中心极限定理与不变原理的关系, 您能用不变原理解决哪些中心极限定理?

7. 在统计的大样本理论中, 不变原理有用吗? 如果有用, 您能用不变原理解决哪些问题?

8. 在定理 4.3 中, 研究的是确定性环境中的经典 Markov 链的不变原理, 对于随机环境中的 Markov 链, 有关它们的不变原理将有什么新面貌? 理论上有重大意义吗?

# 第十章　Lévy 过程和无穷可分律

本章中恒用 d.f. 表分布函数, c.f. 表特征函数, R.V. 表随机变量. 对实变复值函数 $f(t)$, 恒用 $\mathscr{R}(f(t))$ 表其实部.$\mathscr{L}(X)$ 表 R.V.$X$ 的分布.

若 $G(x)=\sigma^2 F(x), F(x)$ 是 d.f., $\sigma^2$ 是正实数, 则称 $G(x)$ 是准分布函数, 用 d$'$.f. 表准分布函数. Lebesgue 可测函数 $g(x)$ 关于 d$'$.f. $G(x)$ 所产生的 L-S 测度 (简记为 $G$) 的积分用

$$\int g(x)dG(x) \quad \left(\text{或} \int g(x)G(dx)\right)$$

表示 (或简记为 $\int gdG$).

前面几章研究的有关的极限定理, 都是随机变量序列 (或随机过程序列 —— 如不变原理) 的极限性质.

本章所研究的极限定理, 是关于随机阵列

$$\{X_{n,k}\}=\begin{Bmatrix} X_{1,1}, \cdots, X_{1,k_1} \\ X_{2,1}, \cdots, X_{2,k_2} \\ \vdots \qquad \vdots \end{Bmatrix}$$

(满足某种性质) 的极限性质.

本章主要研究三个问题:

1. 无穷可分律的特征, 具体到找出其分析表达式;

2. 一个满足某种性质 (下面将要引进的 u.a.n. 性质) 的随机阵列 $\{X_{n,k}\}$, 其

分布

$$\mathscr{L}\left(\sum_{k=1}^{k_n} X_{n,k}\right)$$

弱收敛到某个无穷可分律的充分必要条件;

3. 无穷可分性是 Lévy 过程的特征.

## §1 无穷可分性

**定义 1.1** 称随机阵列

$$\{X_{n,k}\} = \begin{Bmatrix} X_{1,1}, \cdots, X_{1,k_1} \\ X_{2,1}, \cdots, X_{2,k_2} \\ \vdots \qquad \vdots \end{Bmatrix} \tag{1.1}$$

是一个**渐近可忽略体系** (简称 u.a.n. 体系), 如果对每个固定的 $n \geqslant 1, \{X_{n,1}, \cdots; X_{n,k_n}\}$ 是相互独立的随机变量 (即 $\{X_{n,k}\}$ 中每一行内部的随机变量是相互独立的, 但行与行之间未必独立), $\lim\limits_{n\to\infty} k_n = \infty$, 且

$$\lim_{n\to\infty} X_{n,k} = 0, \quad \text{P}. \tag{1.2}$$

对 $k$ 一致成立, 即对任何 $\varepsilon > 0$ 有

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant k_n} P(|X_{n,k}| \geqslant \varepsilon) = 0. \tag{1.3}$$

**定义 1.2** 称 R.V. $X$ 是**无穷可分的** (记之为 i.d.R.V.), 如果对任何正整数 $n \geqslant 1$, 都存在相互独立的具有公共分布的 R.V. $X_1, \cdots, X_n$, 使 $X = X_1 + \cdots + X_n$;

仿之, 称 c.f. $f(t)$ 是**无穷可分的** (记之为 i.d.c.f.), 如果对任何正整数 $n \geqslant 1$, 都存在 c.f. $f_n(t)$, 使 $f(t) = (f_n(t))^n$;

称 d.f.$F(x)$ 是**无穷可分的** (记之为 i.d.d.f.), 如果对任何正整数 $n \geqslant 1$, 都存在 d.f.$F_n(x)$, 使 $F(x) = (F_n * F_n * \cdots * F_n)(x) \overset{\text{def.}}{=\!=} F_n^{n*}(x)$, 其中 $*$ 是卷积符号 $F_n^{n*}$ 表 $F_n(x)$ 产生的 L-S 测度的 $n$ 重卷积.

显然 i.d.R.V. $X$ 的 d.f. 和 c.f. 都是无穷可分的.

**附注 1.1** 如果随机变量阵列 $\{X_{n,k}\}$ 是 u.a.n. 体系, 其相应的特征函数阵列 $\{f_{n,k}(t)\}$ 或相应的分布函数阵列 $\{F_{n,k}(x)\}$ 分别称为 u.a.n.c.f. 体系或 u.a.n.d.f. 体系, 简称 u.a.n. 体系.

**命题 1.1** 下列六条件等价 (即 u.a.n. 体系的定义有六种定义法):

(i) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} P(|X_{nk}|\geqslant\varepsilon)=0$ (一切 $\varepsilon>0$);

(ii) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} \mathscr{R}(1-f_{nk}(t))=0$ (在 $t$ 的任一有限区间上一致成立);

(iii) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} |1-f_{nk}(t)|=0$ (在 $t$ 的任一有限区间上一致成立);

(iv) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} |1-f_{nk}(t)|=0$;

(v) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} \mathscr{R}(1-f_{nk}(t))=0$;

(vi) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n} \int_{\mathbf{R}^1}\dfrac{x^2}{1+x^2}dF_{nk}(x)=0$,

其中 $f_{nk}(t), F_{nk}(x)$ 为 $X_{nk}$ 的 c.f. 和 d.f., $\mathscr{R}(z)$ 表示 $z$ 的实部.

**证** (i)⇒(ii). 考虑 $|t|\leqslant T$. 设 (i) 成立.

$$\begin{aligned}\mathscr{R}(1-f_{nk}(t)) &= \int_{\mathbf{R}^1}(1-\cos xt)dF_{nk}(x)\\ &= \int_{|x|<\varepsilon}(1-\cos xt)dF_{nk}(x)+\int_{|x|\geqslant\varepsilon}(1-\cos xt)dF_{nk}(x)\\ &\leqslant \frac{\varepsilon^2T^2}{2}+\int_{|x|\geqslant\varepsilon}2dF_{nk}(x)\quad (|t|\leqslant T),\end{aligned}$$

由 (i) 得

$$\varlimsup_{n\to\infty}\left(\max_{1\leqslant k\leqslant k_n}\mathscr{R}(1-f_{nk}(t))\right)\leqslant\frac{\varepsilon T^2}{2}\quad (|t|\leqslant T),$$

由 $\varepsilon>0$ 的任意性即得 (ii).

(ii)⇒(iii). 设 (ii) 成立. 由于

$$1-f_{nk}(t)=\mathscr{R}(1-f_{nk}(t))\int_{\mathbf{R}^1} i\sin tx dF_{nk}(x),$$

而

$$\begin{aligned}\left|\int_{\mathbf{R}^1} i\sin tx dF_{nk}(x)\right|^2 &\leqslant \int_{\mathbf{R}^1}\sin^2 tx dF_{nk}(x)\\ &\leqslant \frac{1}{2}\int_{\mathbf{R}^1}(1-\cos 2tx)dF_{nk}(x)\\ &= \frac{1}{2}\mathscr{R}(1-f_{nk}(2t)),\end{aligned}$$

故由 (ii) 可得 (iii).

(iii)⇒(iv)⇒(v) 显然成立.

(v)⇒(vi). 用 [36] 第二章 (4.8) 积分不等式有

$$\max_{1\leqslant k\leqslant k_n}\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x)\leqslant\max_{1\leqslant k\leqslant k_n}M(b)\int_0^b\mathscr{R}(1-f_{nk}(t))dt.$$

再用勒贝格控制收敛定理, 由 (v) 可得 (vi).

(vi)⇒(i). 因为

$$\max_{1\leqslant k\leqslant k_n} P(|X_{nk}|\geqslant\varepsilon)=\max_{1\leqslant k\leqslant k_n}\int_{|x|\geqslant\varepsilon}\frac{x^2}{1+x^2}\frac{1+x^2}{x^2}dF_{nk}(x)$$
$$\leqslant\frac{1+\varepsilon^2}{\varepsilon^2}\max_{1\leqslant k\leqslant k_n}\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x),$$

所以 (vi)⇒(i).

关于 u.a.n. 体系 $\{X_{nk}\}$, 有下列性质:

(i) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n}\mu(X_{nk})=0$ 　($\mu(X_{nk})$ 是 $X_{nk}$ 的中位数);

(ii) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n}\int_{|x|\leqslant\tau}|x|^r dF_{nk}(x)=0$ 　$(r>0,\tau>0)$, 特别地, 若令 $a_{nk}(\tau)=\int_{|x|<\tau}xdF_{nk}(x), a_n(\tau)=\max\limits_{1\leqslant k\leqslant k_n}|a_{nk}(\tau)|$, 则 $\lim\limits_{n\to\infty}a_n(\tau)=0$;

(iii) $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n}|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1|=0$;

(iv) 对于任意的 $b>0$, 存在 $N(b)>0$, 当 $|t|\leqslant b, n\geqslant n\geqslant N(b)$ 时, $\log f_{nk}(t)$ 存在且有限, 还有

$$\log f_{nk}(t)=(f_{nk}(t)-1)+\theta_{nk}(f_{nk}(t)-1)^2,$$
$$\log(e^{-ia_{nk}(\tau)t}f_{nk}(t))=(e^{-ia_{nk}(\tau)t}f_{nk}(t)-1)+\theta'_{nk}(e^{-ia_{nk}(\tau)t}f_{nk}(t)-1)^2,$$

其中 $|\theta_{nk}|\leqslant 1, |\theta'_{nk}|\leqslant 1$;

(v) 对于任意 $b>0,\tau>0$, 存在 $c_1(\tau,b)>0, c_2(\tau,b)>0$, 使

$$c_1(\tau,b)\sup_{|\tau|\leqslant b}|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1|\leqslant\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x+a_{nk}(\tau))$$
$$\leqslant c_2(\tau,b)\int_0^b(1-|f_{nk}(t)|^2)dt$$
$$\leqslant -2c_2(\tau,b)\int_0^b\log|f_{nk}(t)|dt.$$

此处 $F_{nk}(x)$ 和 $f_{nk}(t)$ 分别为 $X_{nk}$ 的 d.f. 和 c.f..

**证**　(i) 因为 $\{X_{nk}\}$ 为 u.a.n. 体系, 故对任何 $\varepsilon>0$, 存在正整数 $N(\varepsilon)$, 当 $n\geqslant N(\varepsilon)$ 时有

$$\max_{1\leqslant k\leqslant k_n}P(|X_{nk}|\geqslant\varepsilon)<\frac{1}{2},$$

从而

$$\max_{1\leqslant k\leqslant k_n}\mu(X_{nk})\in(-\varepsilon,\varepsilon),$$

由 $\varepsilon$ 的任意性知 (i) 成立.

(ii) 任给 $r > 0, \tau > 0$, 有

$$\begin{aligned}
&\max_{1\leqslant k\leqslant k_n}\int_{|x|<\tau}|x|^r dF_{nk}(x)\\
\leqslant &\max_{1\leqslant k\leqslant k_n}\int_{|x|<\varepsilon}|x|^r dF_{nk}(x)+\max_{1\leqslant k\leqslant k_n}\int_{\varepsilon\leqslant|x|<\tau}|x|^r dF_{nk}(x)\\
\leqslant &\varepsilon^r+\tau^r\max_{1\leqslant k\leqslant k_n}\int_{|x|\geqslant\varepsilon}dF_{nk}(x)
\end{aligned}$$

对一切 $0 < \varepsilon < \tau$ 成立, 再用 $\varepsilon > 0$ 可任意小及 $\{X_{nk}\}$ 为 u.a.n. 体系可得 (ii).

(iii) 因为 $\{X_{nk}\}$ 是 u.a.n. 体系, $a_n(\tau) \to 0$, 所以 $\{X_{nk} - a_{nk}(\tau)\}$ 是 u.a.n. 体系, 而 $X_{nk} - a_{nk}(\tau)$ 的特征函数为 $e^{-ia_{nk}(\tau)t}f_{nk}(t)$, 故用 u.a.n. 体系的第 (iv) 个等价条件即得 (iii).

(iv) 由于 $\{X_{nk}\}, \{X_{nk} - a_{nk}(\tau)\}$ 都是 u.a.n. 体系, 所以

$$\max_{1\leqslant k\leqslant k_n}|f_{nk}(t)-1|\to 0,\ \max_{1\leqslant k\leqslant k_n}|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1|\to 0,$$

故 (iv) 成立.

(v) 由于

$$\begin{aligned}
|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1| &= \left|\int_{\mathbf{R}^1}(e^{it(x-a_{nk}(\tau))}-1)dF_{nk}(x)\right|\\
&\leqslant 2\int_{|x|\geqslant\tau}dF_{nk}(x)+\left|\int_{|x|<\tau}(e^{it(x-a_{nk}(\tau))}-1)dF_{nk}(x)\right|\\
&\leqslant 2\int_{|x|\geqslant\tau}dF_{nk}(x)+\left|\int_{|x|<\tau}t(x-a_{nk}(\tau))dF_{nk}(x)\right|\\
&\quad+\left|\int_{|x|<\tau}\frac{t^2(x-a_{nk}(\tau))^2}{2}dF_{nk}(x)\right|\\
&\leqslant 2\int_{|x|\geqslant\tau}dF_{nk}(x)+|t|\left|a_{nk}(\tau)-a_{nk}(\tau)\int_{|x|<\tau}dF_{nk}(x)\right|\\
&\quad+\frac{t^2}{2}\int_{|x|<\tau}[1+(x-a_{nk}(\tau))^2]\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)\\
&\leqslant (2+|ta_{nk}(\tau)|)\int_{|x|\geqslant\tau}dF_{nk}(x)+\frac{t^2}{2}[1+(\tau+|a_{nk}(\tau)|)^2]\\
&\quad\cdot\int_{|x|<\tau}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x).
\end{aligned}$$

但是,

$$\int_{|x|\geqslant\tau} dF_{nk}(x) \leqslant \int_{|x|\geqslant\tau} \frac{1+(\tau-|a_{nk}(\tau)|)^2}{(\tau-|a_{nk}(\tau)|)^2} \cdot \frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2} dF_{nk}(x),$$

由上述二不等式并注意 $\lim\limits_{n\to\infty}\max\limits_{1\leqslant k\leqslant k_n}|a_{nk}(\tau)|=0$ 可知存在 $c(\tau,t)>0$, 使

$$|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1| \leqslant c(\tau,t)\int_{\mathbf{R}^1}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)$$
$$= c'(\tau,t)\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x+a_{nk}(\tau)).$$

所以, 有 $c_1(\tau,b)>0$, 使

$$c_1(\tau,b)\sup_{|t|\leqslant|b|}|e^{-ia_{nk}(\tau)t}f_{nk}(t)-1| \leqslant \int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x+a_{nk}(\tau)).$$

又因为

$$(x-\mu_{nk})^2 \geqslant (x-a_{nk}(\tau))^2 + 2(x-a_{nk}(\tau))(a_{nk}(\tau)-\mu_{nk}),$$

即

$$(x-a_{nk}(\tau))^2 \leqslant (x-\mu_{nk})^2 + 2(x-a_{nk}(\tau))(\mu_{nk}-a_{nk}(\tau)).$$

但是

$$\int_{\mathbf{R}^1}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)$$
$$\leqslant \int_{|x|<\tau}(x-a_{nk}(\tau))^2dF_{nk}(x) + \int_{|x|\geqslant\tau}dF_{nk}(x),$$

而

$$\int_{|x|<\tau}(x-a_{nk}(\tau))^2dF_{nk}(x)$$
$$\leqslant \int_{|x|<\tau}[(x-\mu_{nk})^2+2(x-a_{nk}(\tau))(\mu_{nk}-a_{nk}(\tau))]dF_{nk}(x)$$
$$\leqslant \int_{|x|<\tau}(x-\mu_{nk})^2dF_{nk}(x)$$
$$+2(|\mu_{nk}|+|a_{nk}(\tau)|)|a_{nk}(\tau)|\int_{|x|\geqslant\tau}dF_{nk}(x),$$

所以

$$\begin{aligned}&\int_{\mathbf{R}^1}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)\\ \leqslant&\int_{|x|<\tau}(x-\mu_{nk})^2dF_{nk}(x)\\ &+[2(|\mu_{nk}|+|a_{nk}(\tau)|)|a_{nk}(\tau)|+1]\int_{|x|\geqslant\tau}dF_{nk}(x),\end{aligned}$$

仿前, 可证存在 $D(\tau)>0$, 使

$$\begin{aligned}&\int_{\mathbf{R}^1}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)\\ \leqslant&\ D(\tau)\int_{\mathbf{R}^1}\frac{(x-\mu_{nk})^2}{1+(x-\mu_{nk})^2}dF_{nk}(x)\\ =&\ D(\tau)\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x+\mu_{nk}).\end{aligned}$$

由 [36] 第二章 (4.8) 积分不等式有

$$\int_{\mathbf{R}^1}\frac{x^2}{1+x^2}dF_{nk}(x+\mu_{nk})\leqslant M(b)\int_0^b\{1-|f_{nk}(t)|^2\}dt.$$

总之, 存在 $c_2(\tau,b)>0$, 使

$$\begin{aligned}&\int_{\mathbf{R}^1}\frac{(x-a_{nk}(\tau))^2}{1+(x-a_{nk}(\tau))^2}dF_{nk}(x)\\ \leqslant&\ c_2(\tau,b)\int_0^b\{1-|f_{nk}(t)|^2\}dt\\ \leqslant&\ -2c_2(\tau,b)\int_0^b\log|f_{nk}(t)|dt.\end{aligned}$$

**命题 1.2** 若 $f(t)$ 是 i.d.c.f., 则

(1) $f(t)$ 无处为 0;

(2) 对任何正整数 $n$, 存在 c.f. $f_n(t)$, 使

$$f_n(t)^n=f(t), \text{且} \lim_{n\to\infty}f_n(t)=\perp\quad(\forall t\in\mathbf{R}).$$

**证** (1) 由 $f(t)$ 的无穷可分性知: 对任何正整数 $n$, 存在 c.f. $f_n(t)$, 使 $f_n(t)^n=f(t)$, 所以

$$|f(t)|^{\frac{2}{n}}=|f_n(t)|^2,$$

又因为 $|f(t)|\leqslant 1$, 所以

$$\begin{aligned}\lim_{n\to\infty}|f_n(t)|^2&=\lim_{n\to\infty}|f(t)|^{\frac{2}{n}}\\ &\overset{\text{def.}}{=\!=}g(t)=\begin{cases}1, & \text{当 } f(t)\neq 0;\\ 0, & \text{当 } f(t)=0.\end{cases}\end{aligned}$$

而 $f(t)$ 是连续函数, $f(0)=1$, 所以存在 $\delta>0$, 使得 $|t|\leqslant\delta$ 时, $|f(t)|>\dfrac{1}{2}$, 从而 $|t|\leqslant\delta$ 时, $g(t)=1$, 所以 $g(t)$ 在 0 点连续, 从而 $g(t)$ 也是 c.f., 故 $g(t)\equiv 1(\forall t\in\mathbf{R})$, 所以 $f(t)$ 无处为 0. (1) 证毕.

(2) 若 $f(t)$ 是 i.d.c.f., 由 (1) 知 $f(t)$ 无处为 0, 所以 $\log f(t)$ 存在且有穷. 由 i.d.c.f. 的定义又知对任何正整数 $n$, 存在 c.f.$f_n(t)$, 使 $f_n(t)^n=f(t)$, 故

$$f_n(t)=\exp\left\{\frac{1}{n}\log f(t)\right\}$$

满足 (2) 中的要求. 命题证毕.

## §2 Lévy 过程和 Lévy – Khinchin 公式

**定义 2.1** 称概率空间 $(\Omega,\mathscr{F},P)$ 上的实值随机过程 $X=\{X_t,t\in[0,\infty)\}$ 是 **Lévy 过程**, 如果它具有平稳的相互独立的增量, 即是

(1) 对任何 $h>0, 0\leqslant s<t, (X_{t+h}-X_s)$ 与 $(X_t-X_s)$ 的分布相同;

(2) 对任何正整数 $n\geqslant 2$, 任何 $0\leqslant t_0<t_1<\cdots<t_n, (X_{t_1}-X_{t_0}), (X_{t_2}-X_{t_1}),\cdots,(X_{t_n}-X_{t_{n-1}})$ 相互独立.

对于 Lévy 过程 $\{X_t,t\in[0,\infty)\}$, 不妨设 $X_0\equiv 0$.

**命题 2.1** Lévy 过程 $\{X_t,t\in[0,\infty)\}$ 的每一个随机变量 $X_t$ 都是 i.d.R.V.

证明甚易, 读者可作为习题验证之.

下面我们引进四族特征函数.

$$\mathscr{D}_1=\left\{\begin{array}{l}\text{一切 } \exp\left(i\alpha t+\displaystyle\int_{\mathbf{R}}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\Psi(x),\right.\\ \qquad\text{其中 } \alpha \text{ 是实数}, \Psi(x)=\sigma^2F(x), \sigma^2\geqslant 0, F(x) \text{ 是 d.f.}\end{array}\right\};$$

$$\mathscr{D}_2=\{\text{一切 i.d.c.f.}\};$$

$$\mathscr{D}_3=\left\{f(t)=\prod_{k=1}^{k_n}f_{n,k}(t):\{f_{n,k}(t)\}\text{ 是任一 u.a.n.c.f. 体系}\right\};$$

$$\mathscr{D}_4=\left\{\begin{array}{l}f(t)=\displaystyle\lim_{n\to\infty}e^{iA_nt}\prod_{k=1}^{k_n}f_{n,k}(t):\{A_n\}\text{ 是任一实数列},\\ \qquad\{f_{n,k}(t)\}\text{ 是任一 u.a.n.c.f. 体系.}\end{array}\right\}.$$

注意: $\mathscr{D}_2,\mathscr{D}_3,\mathscr{D}_4$ 是三族特征函数是显然的, 但 $\mathscr{D}_1$ 是一族特征函数却是要证明的.

由于 $\mathscr{D}_1$ 中的函数是由 $\alpha$ 和 $\Psi(x)$ 决定的, 所以 $\mathscr{D}_1$ 中的函数 $f$ 用 $\{\alpha,\Psi(x)\}$ 表示.

$\Psi$ 表示 $\Psi(x)$ 所产生的 L-S 测度.

**命题 2.2** $\mathscr{D}_1$ 中每个 $\{\alpha,\Psi(x)\}$ 都是 c.f..

**证** 注意:

$$\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}$$

是 $x$ 的有界连续函数$\left(\text{在 } x=0 \text{ 的函数值定义为 } \lim\limits_{x\to 0}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}=-\frac{1}{2}t^2\right)$, 而 $\Psi$ 是一个有限的 L-S 测度, 所以若令

$$I_m=\int_{[-A_m,A_m)}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\Psi(x),$$

则有

$$\{\alpha,\Psi(x)\}=\lim_{A_m\to\infty}e^{i\alpha t+I_m}.$$

又因为 $\{\alpha,\Psi(x)\}$ 在 $t=0$ 连续, 所以为证 $\{\alpha,\Psi(x)\}$ 是 c.f., 只需证明 $f_m(t)=e^{I_m}$ 是 c.f.. 今分 $[-A_m,A_m)$ 为

$$-A_m=a_{m0}^{(n)}<a_{m1}^{(n)}<\cdots<a_{mn}^{(n)}=A_m,$$
$$a_{mj}^{(n)}\neq 0,\ \lim_{n\to\infty}\max_{1\leqslant j\leqslant n}(a_{mj}^{(n)}-a_{mj-1}^{(n)})=0,$$

则

$$I_m=\lim_{n\to\infty}\sum_{j=1}^{n}\left(e^{ita_{mj}^{(n)}}-1-\frac{ita_{mj}^{(n)}}{1+(a_{mj}^{(n)})^2}\right)\frac{1+(a_{mj}^{(n)})^2}{(a_{mj}^{(n)})^2}\cdot(\Psi(a_{mj}^{(n)})-\Psi(a_{mj-1}^{(n)})).$$

所以, 存在常数 $C_{mj}^{(n)},\lambda_{mj}^{(n)}\geqslant 0$ 使

$$f_m(t)=e^{I_m}=\lim_{n\to\infty}\prod_{j=1}^{n}e^{-itC_{mj}^{(n)}+\lambda_{mj}^{(n)}(e^{ita_{mj}^{(n)}}-1)},$$

这就说明 $f_m(t)$ 表成了 $n$ 个 c.f. 的乘积的极限, 又因为 $f_m(t)$ 在 $t=0$ 连续, 所以 $f_m(t)$ 是 c.f..

**定理 2.1** $\mathscr{D}_1=\mathscr{D}_2=\mathscr{D}_3=\mathscr{D}_4$.

**证** 任取 $\{\alpha,\Psi(x)\}\in\mathscr{D}_1$, 对任何正整数 $n$,

$$\{\alpha,\Psi(x)\}=\left(\left\{\frac{\alpha}{n},\frac{\Psi(x)}{n}\right\}\right)^n,$$

而由命题 2.2, $\left\{\frac{\alpha}{n},\frac{\Psi(x)}{n}\right\}$ 是 c.f., 所以 $\{\alpha,\Psi(x)\}$ 是 i.d.c.f., 即 $\{\alpha,\Psi(x)\}\in\mathscr{D}_2$.

任取 $f(t) \in \mathscr{D}_2$, 由命题 1.3 (2), 存在 c.f. 列 $\{f_n(t), n \geqslant 1\}$ 使

$$\lim_{n\to\infty} f_n(t) = 1, \quad (f_n(t))^n = f(t) \quad (\forall n \geqslant 1).$$

取 $k_n = n \quad (\forall n \geqslant 1), f_{n,k}(t) = f_n(t) \quad (1 \leqslant k \leqslant k_n = n)$, 则

$$f(t) = \prod_{k=1}^{k_n} f_{n,k}(t).$$

又因为

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant k_n} |f_{n,k}(t) - 1| = \lim_{n\to\infty} |f_n(t) - 1| = 0,$$

所以 $\{f_{n,k}(t)\}$ 是 u.a.n.c.f. 体系, 从而 $f(t) \in \mathscr{D}_3$.

$\mathscr{D}_3 \subset \mathscr{D}_4$ 是显然的.

$\mathscr{D}_4 \subset \mathscr{D}_1$ 留在定理 4.1 中证明.

**定理 2.2** 设 $\{X_t, t \in [0, \infty)\}$ 是 Lévy 过程, 则

(1) 存在 $\{\alpha_1, \Psi_1(x)\} \in \mathscr{D}_1$, 使 $X_1$ 的 c.f. 为 $\boldsymbol{E}(e^{itX_1}) = \{\alpha_1, \Psi_1(x)\}$;

(2) 对任何 $u \geqslant 0, X_u$ 的 c.f. 为 $\boldsymbol{E}(e^{itX_u}) = \{u\alpha_1, u\Psi_1(x)\}$.

**证** (1) 由于 $X_1$ 是 i.d.R.V., 所以 (1) 成立.

(2) 由于 $\{X_t, t \in [0, \infty)\}$ 具有平稳独立增量, 所以对任何正的有理数 $\dfrac{k}{n}$, 有

$$\begin{aligned}
\boldsymbol{E}\left(e^{itX_{\frac{k}{n}}}\right) &= \boldsymbol{E}\left(e^{it\sum\limits_{j=1}^{k}\left(X_{\frac{j}{n}} - X_{\frac{j-1}{n}}\right)}\right) \\
&= \prod_{j=1}^{k} \boldsymbol{E}\left(e^{it\left(X_{\frac{j}{n}} - X_{\frac{j-1}{n}}\right)}\right) \\
&= \boldsymbol{E}(e^{itX_{\frac{1}{n}}})^k = \boldsymbol{E}(e^{itX_1})^{\frac{k}{n}}.
\end{aligned} \tag{2.1}$$

对任何实数 $u \in [0, \infty)$, 总可取正有理数列 $\{r_n\}$, 使 $r_n \to u$. 所以由 (2.1) 立知 (2) 成立.

由定理 2.2 知: 对 Lévy 过程 $\{X_t, t \in [0, \infty)\}, X_1$ 的 c.f. 为

$$f_1(t) = \exp\left(i\alpha_1 t + \int_{\mathbf{R}} \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) \frac{1+x^2}{x^2} d\Psi_1(x)\right), \tag{2.2}$$

(2.2) 式亦可表为

$$f_1(t) = e^{-\Psi(t)}, \tag{2.3}$$

其中

$$\psi(t) = -i\alpha_1 t + \frac{1}{2}t^2 + \int_{\mathbf{R}} \left[1 - e^{itx} + \frac{itx}{1+x^2}\right] \nu(dx), \tag{2.4}$$

$\nu$ 是 $\mathbf{R}$ 上的 Borel 测度, 满足

$$\int_{\mathbf{R}} \frac{x^2}{1+x^2} \nu(dx) < \infty, \quad \nu(\{0\}) = 0. \tag{2.5}$$

由定理 2.2 还知: $X_u$ 的 c.f. 为

$$f_u(t) = \exp\left(i\alpha_1 ut + \int_{\mathbf{R}} u\left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) \frac{1+x^2}{x^2} d\Psi_1(x)\right), \tag{2.6}$$

或

$$f_u(t) = e^{-u\psi(t)}, \quad \psi(t) \text{由 (2.4) 定义}. \tag{2.7}$$

(2.6) 和 (2.7) 都称为 Lévy–Khinchin 公式.

**附注 2.1** 设 $X = \{X_t, t \in [0,\infty)\}$ 是 Lévy 过程, 则 $X$ 的分布由 $X_1$ 的分布唯一决定, 从而 $X_1$ 的 c.f.$\{\alpha_1, \Psi_1(x)\}$ 或 $e^{-\psi(t)}$ 唯一决定, 即是实数 $\alpha_1$ 与 $\Psi_1(x)$ 唯一决定, 或 $\alpha_1$ 与测度 $\nu$ 唯一决定, $\nu$ 称为 Lévy 测度.

**证** 由公式 (2.6)、(2.7), $X_u$ 的分布由 $\alpha_1$ 与 $\Psi_1(x)$ 或 $\alpha_1$ 与 $\nu$ 唯一决定. 再用 $X$ 具有平稳独立增量易证 $X$ 的有限维分布

$$P((X_{u_1}, \cdots, X_{u_n}) \in A_n)$$

由 $X_{u_i} (i = 1, \cdots, n)$ 所唯一决定, 从而 $X_1$ 的分布唯一决定, 故 $X$ 的分布由 $X_1$ 的分布唯一决定.

由此附注看出 Lévy 测度 $\nu$ 是何等重要, 它是刻画 Lévy 过程的概率规律的特征量.

## §3 无穷可分律族的封闭性与连续性

在前一节中, 我们介绍了无穷可分律的分析表示式, 并用标准符号 $\{\alpha, \Psi(x)\}$ 表示. 这一节我们重点证明 $\{\alpha, \Psi(x)\}$ 的表示法唯一, 且在某种意义下连续, 即是

(1) 若 $\{\alpha_1, \Psi_1(x)\} = \{\alpha_2, \Psi_2(x)\} \in \mathscr{D}_1$, 则 $\alpha_1 = \alpha_2, \Psi_1(x) = \Psi_2(x)$.

(2) 若 $\{\alpha_n, \Psi_n(x)\} \in \mathscr{D}_1 (n \geqslant 1), \alpha_n \to \alpha, \Psi_n \xrightarrow{\text{w}} \Psi, \{\alpha, \Psi(x)\} \in \mathscr{D}_1$, 则 $\{\alpha_n, \Psi_n(x)\} \to \{\alpha, \Psi(x)\}$.

**定理 3.1 (唯一性定理)** *$\{\alpha, \Psi(x)\}$ 唯一地决定了 $\alpha$ 及 $\Psi(x)$.*

**证** 令 $\{\alpha, \Psi(x)\} = e^{\psi(t)}$, 作

$$\begin{aligned}\varphi(t) &= \psi(t) - \int_0^1 \frac{\psi(t+h)+\psi(t-h)}{2} dh \\ &= \int_0^1 \left[\int_{\mathbf{R}^1} (1-\cos hx) e^{itx} \frac{1+x^2}{x^2} d\Psi(x)\right] dh \\ &= \int_{\mathbf{R}^1} e^{itx} \left(1 - \frac{\sin x}{x}\right) \frac{1+x^2}{x^2} d\Psi(x).\end{aligned}$$

若令

$$\Phi(x) = \int_{(-\infty,x)} \left(1 - \frac{\sin y}{y}\right) \frac{1+y^2}{y^2} d\Psi(y),$$

则 $\Phi(x)$ 等于一个分布函数乘以非负常数 $\sigma^2$, 且

$$\varphi(t) = \int_{\mathbf{R}^1} e^{itx} d\Phi(x).$$

由傅氏变换的反演公式知 $\varphi(t)$ 唯一决定了 $\Phi(x)$, 所以 $e^{\psi(t)}$ 唯一决定了

$$\Psi(x) = \int_{(-\infty,x)} \frac{1}{\left(1 - \frac{\sin y}{y}\right) \frac{1+y^2}{y^2}} d\Phi(y).$$

又因为 $\alpha$ 由 $\Psi(x)$ 及 $\{\alpha, \Psi(x)\}$ 所唯一决定, 故 $\{\alpha, \Psi(x)\}$ 也唯一决定了 $\alpha$. 定理证毕.

**定理 3.2 (封闭性及连续性)** (i) $\alpha_n \to \alpha, \Psi_n(x) \xrightarrow{\text{w}} \Psi(x) \Rightarrow \{\alpha_n, \Psi_n(x)\} \to \{\alpha, \Psi(x)\}$;

(ii) 若 $\{\alpha_n, \Psi_n(x)\} \to f, f$ 是 c.f., 则 $\alpha_n \to \alpha, \Psi_n(x) \xrightarrow{\text{w}} \Psi(x)$, 且 $f = \{\alpha, \Psi(x)\}$.

**证** (i) 是第三章定理 5.9 的直接推论.

(ii) 因为 $\{\alpha_n, \Psi_n(x)\} \to f, f$ 是 c.f., 所以

$\{\alpha_n, \Psi_n(x)\} \to f(t)$ (在 $|t| \leqslant T$ 上一致成立),

因此

$$\begin{aligned}|\{\alpha_n, \Psi_n(x)\}| &= \left| e^{\int_{\mathbf{R}^1} \left(e^{itx} - 1 - \frac{itx}{1+x^2}\right) \frac{1+x^2}{x^2} d\Psi_n(x)} \right| \\ &= \left| e^{\int_{\mathbf{R}^1} (\cos tx - 1) \frac{1+x^2}{x^2} d\Psi_n(x)} \right| \to |f(t)|,\end{aligned}$$

在 $|t| \leqslant T$ 上一致成立.

取 $\delta > 0$, 使 $|f(t)| > \dfrac{1}{2}$ (当 $|t| \leqslant \delta$ 时), 不妨令 $\delta < T$, 则

$$\lim_{n\to\infty} \int_{\mathbf{R}^1} (1 - \cos tx) \frac{1+x^2}{x^2} d\Psi_n(x) = -\log|f(t)| \leqslant c_1$$

在 $|t|\leqslant\delta$ 上一致成立. 因此存在正数 $c'$ 使

$$\frac{1}{\delta}\int_0^\delta\left[\int_{\mathbf{R}^1}(1-\cos tx)\frac{1+x^2}{x^2}d\Psi_n(x)\right]dt\leqslant c'\quad(n\geqslant1),$$

此即

$$\int_{\mathbf{R}^1}\left(1-\frac{\sin\delta x}{\delta x}\right)\frac{1+x^2}{x^2}d\Psi_n(x)\leqslant c'\quad(n\geqslant1).$$

由于存在正数 $A$ 和 $B$, 使

$$0<A\leqslant\left(1-\frac{\sin\delta x}{\delta x}\right)\frac{1+x^2}{x^2}\leqslant B<\infty\quad(\text{对一切 }x\in\mathbf{R}^1),$$

所以 $\sup\limits_{x\in\mathbf{R}^1}\sup\limits_{n\geqslant1}\Psi_n(x)\leqslant c$. 但是

$$(1-\cos tx)\frac{1+x^2}{x^2}\leqslant\left|\frac{1-\cos tx}{x^2}\right|+|1-\cos tx|\leqslant\frac{t^2}{2}+2,$$

所以

$$\int_{\mathbf{R}^1}(1-\cos tx)\frac{1+x^2}{x^2}d\Psi_n(x)\leqslant\left(2+\frac{t^2}{2}\right)c\quad(n\geqslant1).$$

因此

$$|f(t)|\geqslant e^{-\left(2+\frac{t^2}{2}\right)c}>0\quad(t\in\mathbf{R}^1)$$

从而 $\log|f(t)|$ 存在且有穷 $(t\in\mathbf{R}^1)$. 令

$$\psi_n(t)=i\alpha_nt+\int_{\mathbf{R}^1}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\Psi_n(x),$$

则

$$\lim_{n\to\infty}\psi_n(t)=\log f(t)\quad\text{在 }|t|\leqslant T\text{ 上一致成立.}$$

若令

$$\psi(t)=\log f(t),\quad q(y)=\left(1-\frac{\sin y}{y}\right)\frac{1+y^2}{y^2},$$

$$\Phi_n(x)=\int_{(-\infty,x)}q(y)d\Psi_n(y),$$

则由 $\Psi_n(x)$ 的一致有界性知 $\psi_n(t)$ 在 $t\in[0,1]$ 上一致有界, 故由控制收敛定理有

$$\begin{aligned}&\lim_{n\to\infty}\int_{\mathbf{R}^1}e^{itx}d\Phi_n(x)\\&=\lim_{n\to\infty}\left(\psi_n(t)-\int_0^1\frac{\psi_n(t+h)+\psi_n(t-h)}{2}dh\right)\\&=\psi(t)-\int_0^1\frac{\psi(t+h)+\psi(t-h)}{2}dh.\end{aligned}$$

上式右端在 $t=0$ 连续, 又因为 $\Phi_n(x)$ 可表为一个分布函数乘以非负常数, 且 $\sup\limits_{n\geqslant 1}\sup\limits_{x\in\mathbf{R}^1}\Phi_n(x)\leqslant D<\infty$, 因此存在 $\Phi(x)$, 它等于一个分布函数乘以非负常数, 并有

$$\Phi_n(x)\xrightarrow{\text{w}}\Phi(x)\quad(n\to\infty),$$

从而存在 $\Psi(x)$, 它亦为某一分布函数乘以非负常数, 并有

$$\Psi_n(x)\xrightarrow{\text{w}}\Psi(x)\quad(n\to\infty).$$

所以

$$\begin{aligned}&\lim_{n\to\infty}\int_{\mathbf{R}^1}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\Psi_n(x)\\=&\int_{\mathbf{R}^1}\left(e^{itx}-1-\frac{itx}{1+x^2}\right)\frac{1+x^2}{x^2}d\Psi(x).\end{aligned}$$

又因为

$$\lim_{n\to\infty}\psi_n(t)=\psi(t),$$

把上述两式相减, 即发现存在 $\alpha$, 使

$$\lim_{n\to\infty}\alpha_n=\alpha.$$

再用本定理的 (i) 知

$$\lim_{n\to\infty}\{\alpha_n,\Psi(x)\}=\lim_{n\to\infty}e^{\psi_n(t)}=\{\alpha,\Psi(x)\}.$$

但前面已证

$$\lim_{n\to\infty}e^{\psi_n(t)}=f(t),$$

所以 $f(t)=\{\alpha,\Psi(x)\}$. 定理证毕.

**命题 3.1** $\mathscr{D}_1$ 重合于: 有限个 Poisson 型的特征函数 $f_k(t)=\exp(i\alpha_k t+\lambda_k(e^{i\alpha_k t}-1))$ 之积的极限特征函数族.

**证** 在命题 2.2 中已证前者含于后者. 为证命题 3.1, 只证后者含于前者. 任取

$$f(t)=e^{i\alpha t+\lambda(e^{i\alpha t}-1)},$$

作 $\Psi(x)$ 如下: $\Psi$ 的测度集中的 $\alpha$ 点, 且 $\Psi(\{\alpha\})=\lambda\alpha^2/(1+\alpha^2)$, 则 $f(t)=\left\{\alpha+\dfrac{\lambda\alpha}{1+\alpha},\Psi(x)\right\}\in\mathscr{D}_1$, 再注意 $\mathscr{D}_1$ 对函数的乘法运算和极限运算是封闭的, 则命题得证.

**命题 3.2** $f(t) \in \mathscr{D}_1 \Longrightarrow e^{i\alpha t} f(bt) \in \mathscr{D}_1$.

**证** 显然.

例如, 下列诸特征函数是无穷可分的.

(1) 退化分布的特征函数 $e^{i\alpha t}$;

(2) 正态分布的特征函数 $e^{i\alpha t - \frac{1}{2}\sigma^2 t^2}$;

(3) Poisson 分布的特征函数 $e^{\lambda(e^{it}-1)}$;

(4) 有限个或可数个 Poisson 型的特征函数之积

$$f(t) = e^{\sum\limits_j [i\alpha_j^t + \lambda_j (e^{i\alpha_j t} - 1)]}$$

$\left(\text{其中 } \sum_j \alpha_j \text{ 收敛}, \alpha_j \neq 0, \lambda_j > 0, \sum_j \lambda_j < \infty\right)$;

(5) Cauchy 分布 $F(x) = \dfrac{a}{\pi} \int_{-\infty}^{x} \dfrac{dy}{a^2 + (y-\alpha)^2}$ 的特征函数

$$f(t) = e^{i\alpha t - \alpha|t|};$$

(6) Pearson 第三型分布函数的特征函数

$$f(t) = (1 - it)^c \quad (c > 0).$$

**附注 3.1** 无处为 0 的特征函数不一定是无穷可分的. 例如 R.V.X 定义如下: $X$ 仅取 $-1, 0, 1$ 三个值, 对应的概率为 $\dfrac{1}{8}, \dfrac{3}{4}, \dfrac{1}{8}$. 其特征函数为 $f(t) = \dfrac{1}{8}e^{-it} + \dfrac{3}{4} + \dfrac{1}{8}e^{it} = \dfrac{3 + \cos t}{4} > 0$, 但 $f(t)$ 不是无穷可分的. 事实上, 若 $X$ 与 $X_1 + X_2$ 的分布函数相同, $X_1$ 与 $X_2$ 独立同分布, 则 $X_1$ 最多只能在二个不同点上有正概率, 不妨设 $a_1 < a_2, P(X_1 = a_1) = p, P(X_1 = a_2) = 1 - p$, 于是 $P(X = 2a_1) = p^2, P(X = a_1 + a_2) = p(1-p), P(X = 2a_2) = (1-p)^2$. 由 $a_1 < a_2$ 得 $2a_1 = -1, a_1 + a_2 = 0, 2a_2 = 1, p^2 = \dfrac{1}{8}, p(1-p) = \dfrac{3}{4}, (1-p)^2 = \dfrac{1}{8}$, 而此为不可能.

## §4 u.a.n. 体系的极限特征函数族

在这一节中, 我们要补证完定理 2.1, 即是要补证 $\mathscr{D}_4 \subset \mathscr{D}_1$, 亦即要证明 u.a.n. 体系的极限特征函数族, 就是无穷可分特征函数族 $\mathscr{D}_1$.

**引理 4.1** 设 $\{f_{nk}(t)\}$ 是 u.a.n. 体系, 且

$$\sum_{k=1}^{k_n} \int_{\mathbf{R}^1} \frac{x^2}{1+x^2} dF_{nk}(x + a_{nk}(\tau)) \leqslant c < \infty \quad (c \text{ 与 } n \text{ 无关}),$$

$$e^{-iA_n t}\prod_{k=1}^{k_n} f_{nk}(t)=\theta_n(t)\{\alpha_n,\Psi_n(x)\},$$

其中 $F_{nk}(x)$ 是 $f_{nk}(t)$ 之 d.f., $a_{nk}(\tau)=\int_{|x|<\tau} x dF_{nk}(x)$ 如 §3 所定义,$\theta_n(t)\to 1, A_n$ 是实数,

$$c_n=-A_n+\sum_{k=1}^{k_n} a_{nk}(\tau)+\sum_{k=1}^{k_n}\int_{\mathbf{R}^1}\frac{x}{1+x^2}dF_{nk}(x+a_{nk}(\tau)),$$

$$\Psi_n(x)=\sum_{k=1}^{k_n}\int_{(-\infty,x)}\frac{y^2}{1+y^2}dF_{nk}(y+a_{nk}(\tau)).$$

以后在不会混淆的情况下, 简记 $a_{nk}(\tau)$ 为 $a_{nk}$, $\prod_{k=1}^{k_n}\left(\text{或}\ \sum_{k=1}^{k_n}\right)$ 为 $\prod_k\left(\text{或}\ \sum_k\right)$, $\max_{1\leqslant k\leqslant k_n}$ 为 $\max_k$.

**证** 令 $\tilde{f}_{nk}(t)=e^{-ita_{nk}}f_{nk}(t)$, 任取 $t_0\in\mathbf{R}^1$, 再令 $z_{nk}=\tilde{f}_{nk}(t_0),\widetilde{A}_n=A_n-\sum_k a_{nk}, W_n=e^{-iA_n t_0}\prod_k f_{nk}(t_0)$, 则

$$\begin{aligned}W_n&=e^{-i\widetilde{A}_n t_0}\prod_k z_{nk}=e^{-i\widetilde{A}_n t_0}\cdot e^{\sum_k \log z_{nk}}\\&=e^{-i\widetilde{A}_n t_0}e^{\sum_k((z_{nk}-1)-\frac{1}{2}(z_{nk}-1)^2+\cdots)}\end{aligned}$$

而

$$\begin{aligned}|z_{nk}-1|&=|f_{nk}(t_0)(e^{-ia_{nk}t_0}-1)+(f_{nk}(t_0)-1)|\\&\leqslant|e^{-ia_{nk}t_0}-1|+|f_{nk}(t_0)-1|,\end{aligned}$$

由 $\{f_{nk}\}$ 是 u.a.n. 体系知

$$\lim_{n\to\infty}\max_k|a_{nk}|=0,\quad \lim_{n\to\infty}\max_k|f_{nk}(t)-1|=0,$$

所以, 若令 $\xi_n=\max_k|z_{nk}-1|$, 则有 $\xi_n=o(1)\ (n\to\infty)$. 因此, 由 u.a.n. 之性质 (v) 有

$$\begin{aligned}\sum_k|z_{nk}-1|&\leqslant A(\tau)\sum_k\int_{\mathbf{R}}\frac{x^2}{1+x^2}dF_{nk}(x+a_{nk})\\&\leqslant c<\infty\quad(n\geqslant 1).\end{aligned}$$

$$\sum_k|z_{nk}-1|^2\leqslant\xi_n\sum_k|z_{nk}-1|=o(1)\quad(n\to\infty).$$

因此, 取

$$\theta_n(t_0) = e^{-\sum\limits_k [\frac{1}{2}(z_{nk}-1)^2 - \cdots]}$$

则有 $\lim\limits_{n\to\infty} \theta_n(t_0) = 1$, 且

$$\begin{aligned} W_n &= \theta_n(t_0) e^{-i\widetilde{A}_n t_0} \bullet e^{\sum\limits_k (z_{nk}-1)} \\ &= \theta_n(t_0) e^{-i\widetilde{A}_n t_0} \bullet e^{\sum\limits_k \int_{\mathbf{R}^1} (e^{it_0 x}-1) dF_{nk}(x+a_{nk})} \\ &= \theta_n(t_0) e^{i\alpha_n t_0 + \int_{\mathbf{R}^1} \left(e^{it_0 x} - 1 - \frac{it_0 x}{1+x^2}\right) \frac{1+x^2}{x^2} d\Psi_n(x)} \end{aligned}$$

**定理 4.1 (广义中心极限定理)** 设 $\{f_{nk}\}$ 为 u.a.n 体系, $\{A_n\}$ 为一串实数, 若

$$\lim_{n\to\infty} e^{-iA_n t} \prod_k f_{nk}(t) = f(t),$$

$f(t)$ 是 c.f., 则 $f = \{\alpha, \Psi(x)\} \in \mathscr{D}_1$, $\mathscr{D}_1$ 是 §2 中所定义的无穷可分特征函数族.

**证** 由 u.a.n. 体系性质 (v) 有

$$\begin{aligned} &\sum_k \int_{\mathbf{R}^1} \frac{x^2}{1+x^2} dF_{nk}(x+a_{nk}) \\ &\leqslant -2c_2(\tau,\delta_0) \int_0^{\delta_0} \sum_k \log|f_{nk}(t)| dt. \end{aligned} \tag{4.1}$$

而

$$\lim_{n\to\infty} \prod_k |f_{nk}(t)|^2 = |f(t)|^2, \tag{4.2}$$

又因为特征函数序称收敛到特征函数在有限区间上总是一致的, 且 $|f(t)|^2$ 连续, $f(0)=1$, 所以可以取适当小的 $\delta_0 > 0$, 使

$$0 \leqslant -\log \prod_k |f_{nk}(t)|^2 \leqslant M \quad (\text{一切 } |t| \leqslant \delta_0, n \geqslant 1),$$

也就是

$$0 \leqslant -2\sum_k \log|f_{nk}(t)| \leqslant M \quad (\text{一切 } |t| \leqslant \delta_0, n \geqslant 1), \tag{4.3}$$

由 (4.1) 和 (4.3) 得知引理 4.1 的条件成立, 故由引理 4.1 得

$$e^{-iA_n t} \prod_k f_{nk}(t) = \theta_n(t)\{\alpha_n, \Psi_n(x)\}, \quad \theta_n(t) \to 1.$$

而 $\lim\limits_{n\to\infty} \theta_n(t)\{\alpha_n, \Psi_n(x)\} = f(t)$, 所以

$$f(t) = \lim_{n\to\infty} \{\alpha_n, \Psi_n(x)\}.$$

而 $\{\alpha_n, \Psi_n(x)\} \in \mathscr{D}_1$, 所以由 §3 定理 3.2 $\mathscr{D}_1$ 的封闭性知 $f(t) \in \mathscr{D}_1$.

本定理完成了定理 2.1 中未证明的部分: $\mathscr{D}_4 \subset \mathscr{D}_1$

下面我们研究 $e^{-iA_n t}\prod\limits_k f_{nk}(t) \to \{\alpha, \Psi(x)\}$ 的充要条件.

## §5 收敛到无穷可分律的充分必要条件

**定理 5.1** 若 $\{f_{nk}(t)\}$ 为 u.a.n. 体系, 则

$$\lim_{n\to\infty} e^{-iA_n t}\prod_k f_{nk}(t) = \{\alpha, \Psi(x)\}$$

的充要条件是

$$\begin{cases} \Psi_n(x) \xrightarrow{\text{w}} \Psi(x), & (\text{c}_1)\\ \alpha_n \to \alpha \quad (\text{对一个或一切 } \tau > 0), & (\text{c}_2)\end{cases}$$

其中

$$\Psi_n(x) = \sum_k \int_{(-\infty,x)} \frac{y^2}{1+y^2} dF_{nk}(y + a_{nk}),$$

$$\alpha_n = -A_n + \sum_k a_{nk} + \sum_k \int_{\mathbf{R}^1} \frac{x^2}{1+x^2} dF_{nk}(x + a_{nk}).$$

注意: 由于 $a_{nk} = a_{nk}(\tau)$, 所以 $\Psi_n(x)$ 与 $\alpha_n$ 都和 $\tau$ 有关.

**证** 必要性. 仿定理 4.1, 可证

$$\sum_k \int_{\mathbf{R}^1} \frac{x^2}{1+x^2} dF_{nk}(x + a_{nk}) \leqslant c < \infty \quad (\text{与 } n \text{ 无关}).$$

故由引理 4.1 有

$$e^{-iA_n t}\prod_k f_{nk}(t) = \theta_n(t)\{\alpha_n, \Psi_n(x)\}, \quad \theta_n(t) \to 1.$$

而

$$\lim_{n\to\infty} e^{-iA_n t}\prod_k f_{nk}(t) = \{\alpha, \Psi(x)\},$$

所以 $\lim\limits_{n\to\infty}\{\alpha_n, \Psi_n(x)\} = \{\alpha, \Psi(x)\}$. 用定理 3.2 得: $\alpha_n \to \alpha, \Psi_n(x) \xrightarrow{\text{w}} \Psi(x)$.

充分性. 若 $\{f_{nk}(t)\}$ 是 u.a.n. 体系, 且

$$\alpha_n \to \alpha, \quad \Psi_n(x) \xrightarrow{\text{w}} \Psi(x).$$

则由 $\lim\limits_{n\to\infty}\Psi_n(\infty) = \Psi(\infty) < \infty$ 有

$$\sum_k \int_{\mathbf{R}^1} \frac{x^2}{1+x^2} dF_{nk}(x + a_{nk})$$
$$= \Psi_n(\infty) \leqslant c < \infty \quad (\text{与 } n \text{ 无关}),$$

此即引理 4.1 的条件全部满足, 所以

$$e^{-iA_nt}\prod_k f_{nk}(t)=\theta_n(t)\{\alpha_n,\Psi_n(x)\},\quad \theta_n(t)\to 1.$$

而由定理 3.2 有

$$\{\alpha_n,\Psi_n(x)\}\to\{\alpha,\Psi(x)\}.$$

所以

$$\lim_{n\to\infty}e^{-iA_nt}\prod_k f_{nk}(t)=\{\alpha,\Psi(x)\}.$$

定理证毕.

由于

$$\Psi_n(x)\xrightarrow{\mathrm{w}}\Psi(x),\quad \alpha_n\to\alpha$$

的概率意义不太明确, 下面我们再给出一个充分必要条件. 为此, 需要作一些准备工作.

**两个不等式**　设 $\{F_{nk}(x)\}$ 为 u.a.n. 体系.

$$\text{(i)}\qquad \begin{aligned}&\sum_k\int_{|y|<x}y^2dF_{nk}(y+a_{nk}(\tau))\\ &=\sum_k\left\{\int_{|y|<x}y^2dF_{nk}(y)-\left(\int_{|y|<x}ydF_{nk}(y)\right)^2\right\}+\Delta_n,\end{aligned}\tag{5.1}$$

其中不妨令 $0<x<\tau$, 且

$$|\Delta_n|\leqslant\varepsilon\sum_k\int_{|y|\geqslant x}dF_{nk}(y)+(x+2\varepsilon)^2\sum_k\int_{x-\varepsilon\leqslant|y|<x+\varepsilon}dF_{nk}(y)\tag{5.2}$$

(只要 $n\geqslant N(\varepsilon,x,\tau)$).

证

$$
\begin{aligned}
&\sum_k \int_{|y|<x} y^2 dF_{nk}(y+a_{nk}(\tau)) \\
= &\sum_k \int_{|y-a_{nk}(\tau)|<x} (y-a_{nk}(\tau))^2 dF_{nk}(y) \\
\overset{\text{记作}}{=\!=} &\sum_k \int_{|y|<x} (y-a_{nk}(\tau))^2 dF_{nk}(y) + \Delta_n^{(1)} \\
\overset{\text{记作}}{=\!=} &\sum_k \int_{|y|<x} (y-a_{nk}(x))^2 dF_{nk}(y) + \Delta_n^{(2)} + \Delta_n^{(1)} \\
= &\sum_k \left[\int_{|y|<x} y^2 dF_{nk}(y) - 2a_{nk}(x)^2 + \int_{|y|<x} a_{nk}(x)^2 dF_{nk}(y)\right] + \Delta_n^{(2)} + \Delta_n^{(1)} \\
\overset{\text{记作}}{=\!=} &\sum_k \left[\int_{|y|<x} y^2 dF_{nk}(y) - \left(\int_{|y|<x} y dF_{nk}(y)\right)^2\right] \\
&+\Delta_n^{(3)} + \Delta_n^{(2)} + \Delta_n^{(1)}.
\end{aligned} \tag{5.3}
$$

而

$$
\begin{aligned}
|\Delta_n^{(3)}| &\leqslant \sum_k \left| a_{nk}(x)^2 - a_{nk}(x)^2 \int_{|y|<x} dF_{nk}(y) \right| \\
&= \sum_k a_{nk}(x)^2 \int_{|y|\geqslant x} dF_{nk}(y) \\
&\leqslant a_n(x)^2 \sum_k \int_{|y|\geqslant x} dF_{nk}(y) \quad \left(a_n = \max_k |a_{nk}|\right), \\
\Delta_n^{(2)} &= \sum_k \int_{|y|<x} (a_{nk}(x) - a_{nk}(\tau))(2y - a_{nk}(x) - a_{nk}(\tau)) dF_{nk}(y) \\
&= \sum_k (a_{nk}(x) - a_{nk}(\tau)) \int_{|y|<x} [2(y - a_{nk}(x)) + (a_{nk}(x) - a_{nk}(\tau))] dF_{nk}(y) \\
&= \sum_k \left[ 2(a_{nk}(x) - a_{nk}(\tau)) a_{nk}(x) \int_{|y|\geqslant x} dF_{nk}(y) \right. \\
&\quad \left. + (a_{nk}(x) - a_{nk}(\tau))^2 \int_{|y|<x} dF_{nk}(y) \right],
\end{aligned}
$$

所以若注意 $a_n = \max\limits_k |a_{nk}|$ 则得

$$\begin{aligned}
|\Delta_n^{(2)}| &\leqslant 2(a_n(x)+a_n(\tau))a_n(x)\sum_k \int_{|y|\geqslant x} dF_{nk}(y) \\
&\quad + \sum_k \left(\int_{x\leqslant|y|<\tau} y dF_{nk}(y)\right)^2 \\
&\leqslant 2(a_n(x)+a_n(\tau))a_n(x)\sum_k \int_{|y|\geqslant x} dF_{nk}(y) \\
&\quad + \sum_k \tau^2 \left(\int_{|y|\geqslant x} dF_{nk}(y)\right)^2 \\
&\leqslant 2(a_n(x)+a_n(\tau))a_n(x)\sum_k \int_{|y|\geqslant x} dF_{nk}(y) \\
&\quad + \tau^2 \left(\max_k \int_{|y|\geqslant x} dF_{nk}(y)\right)\sum_k \int_{|y|\geqslant x} dF_{nk}(y) \\
&\leqslant \left[2(a_n(x)+a_n(\tau))a_n(x) + \tau^2 \max_k \int_{|y|\geqslant x} dF_{nk}(y)\right] \\
&\quad \cdot \left[\sum_k \int_{|y|\geqslant x} dF_{nk}(y)\right].
\end{aligned}$$

因为在 u.a.n. 条件下有

$$\lim_{n\to\infty} a_n(x) = 0, \quad \lim_{n\to\infty} \max_k \int_{|y|\geqslant x} dF_{nk}(y) = 0,$$

所以, 对任何 $\varepsilon > 0$, 都存在 $N(\varepsilon, x, \tau)$, 使 $n \geqslant N(\varepsilon, x, \tau)$ 时有

$$|\Delta_n^{(3)}| + |\Delta_n^{(2)}| \leqslant \varepsilon \sum_k \int_{|y|\geqslant x} dF_{nk}(y). \tag{5.4}$$

又

$$\Delta_n^{(1)} = \sum_k \left[\int_{|y-a_{nk}(\tau)|<x} (y-a_{nk}(\tau))^2 dF_{nk}(y) - \int_{|y|<x} (y-a_{nk}(\tau))^2 dF_{nk}(y)\right],$$

所以

$$|\Delta_n^{(1)}| \leqslant \sum_k \int_V (y-a_{nk}(\tau))^2 dF_{nk}(y),$$

其中 $V = \{y | x - |a_{nk}(\tau)| \leqslant |y| \leqslant x + |a_{nk}(\tau)|\}$. 因此, 当 $n$ 充分大后有

$$\begin{aligned}
|\Delta_n^{(1)}| &\leqslant \sum_k \int_{x-\varepsilon\leqslant|y|\leqslant x+\varepsilon} \{|y|+\varepsilon\}^2 dF_{nk}(y) \\
&\leqslant (x+2\varepsilon)^2 \sum_k \int_{x-\varepsilon\leqslant|y|\leqslant x+\varepsilon} dF_{nk}(y).
\end{aligned} \tag{5.5}$$

由 (5.1)~(5.3) 即得证本不等式.

(ii)

$$\left|\sum_k \int_{|y|<\tau} y dF_{nk}(y + a_{nk}(\tau))\right|$$
$$\leqslant (\tau + 2\varepsilon) \sum_k \int_{\tau-\varepsilon\leqslant|y|\leqslant\tau+\varepsilon} dF_{nk}(y) + \varepsilon \sum_k \int_{|y|\geqslant\tau} dF_{nk}(y) \tag{5.6}$$

(只要 $n \geqslant N(\varepsilon, \tau)$).

**证** $\left|\sum_k \int_{|y|<\tau} y dF_{nk}(y + a_{nk}\tau)\right|$

$$= \left|\sum_k \int_{|y-a_{nk}(\tau)|<\tau} (y - a_{nk}(\tau)) dF_{nk}(y)\right|$$
$$\leqslant \left|\sum_k \left(\int_{|y-a_{nk}(\tau)|<\tau} (y - a_{nk}(\tau)) dF_{nk}(y) - \int_{|y|<\tau} (y - a_{nk}(\tau)) dF_{nk}(y)\right)\right|$$
$$+ \left|\sum_k \int_{|y|<\tau} (y - a_{nk}(\tau)) dF_{nk}(y)\right|$$
$$\leqslant \sum_k \int_U (|y| + |a_{nk}(\tau)|) dF_{nk}(y)$$
$$+ \left|\sum_k a_{nk}(\tau) \int_{|y|\geqslant\tau} dF_{nk}(y)\right|, \tag{5.7}$$

其中 $U = \{y | \tau - |a_{nk}(\tau)| \leqslant |y| \leqslant \tau + |a_{nk}(\tau)|\}$. 由于 $\{F_{nk}(x)\}$ 是 u.a.n. 体系, 所以

$$\lim_{n\to\infty} \max_k |a_{nk}(\tau)| = 0,$$

因此, 当 $n \geqslant N(\varepsilon, \tau)$ 时有

$$\max_k |a_{nk}(\tau)| \leqslant \varepsilon. \tag{5.8}$$

以 (5.8) 代入 (5.7) 得

$$\left|\sum_k \int_{|y-a_{nk}(\tau)|<\tau} (y - a_{nk}(\tau)) dF_{nk}(y)\right|$$
$$\leqslant \sum_k \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} (|y| + \varepsilon) dF_{nk}(y) + \varepsilon \sum_k \int_{|y|\geqslant\tau} dF_{nk}(y)$$
$$\leqslant (\tau + 2\varepsilon) \sum_k \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} dF_{nk}(y) + \varepsilon \sum_k \int_{|y|\geqslant\tau} dF_{nk}(y). \tag{5.9}$$

**引理 5.1**　*令 $\Psi_1,\Psi_2$ 为 $\mathscr{B}^1$ 上的两个有限测度, 若*

(i) $\int_{(-\infty,x)}\dfrac{1+y^2}{y^2}d\Psi_1(y)=\int_{(-\infty,x)}\dfrac{1+y^2}{y^2}d\Psi_2(y)$,

$$(x<0, x\in C(\Psi_1)\bigcap C(\Psi_2));$$

(ii) $\int_{[x,\infty)}\dfrac{1+y^2}{y^2}d\Psi_1(y)=\int_{[x,\infty)}\dfrac{1+y^2}{y^2}d\Psi_2(y)$,

$$(x>0, x\in (\Psi_1)\bigcap(\Psi_2));$$

(iii) $\Psi_1(\{0\})=\Psi_2(\{0\})$,

*则* $\Psi_1\equiv\Psi_2$.

**证**　(1) 先考虑负半轴. 令

$$\Phi_i(x)=\int_{(-\infty,x)}\frac{1+y^2}{y^2}d\Psi_i(y)\quad (x<0), i=1,2,$$

则由 (i) 知 $\Phi_1(x)=\Phi_2(x), x<0, x\in C(\Psi_1)\bigcap C(\Psi_2)$, $C(\Psi)$ 表示 $\Psi$ 的**连续点集**. 而 $C(\Psi_1)\bigcap C(\Psi_2)$ 在 $\mathbf{R}^1$ 中处处稠密, 且 $\Phi_i(x)$ 是 $x$ 的单调非降左连续的函数, 所以

$$\Phi_1(x)\equiv\Phi_2(x),\quad x<0.$$

因此任取 $a<b<0$, 有

$$\Psi_1([a,b))=\int_{[a,b)}\frac{x^2}{1+x^2}d\Phi_1(x)$$
$$=\int_{[a,b)}\frac{x^2}{1+x^2}d\Phi_2(x)=\Psi_2([a,b)).$$

而一切区间 (包括空集) 成半环, 故由测度扩张的唯一性定理得

$$\Psi_1(A)=\Psi_2(A),\quad A\in\mathscr{B}^1(-\infty,0),$$

$\mathscr{B}^1(-\infty,0)$ 表示负半轴上的全体 Borel 集合.

(2) 仿 (1) 可证: $\Psi_1(A)=\Psi_2(A), A\in\mathscr{B}^1(0,\infty)$.

最后, 由 (iii) 有 $\Psi_1(\{0\})=\Psi(\{0\})$, 所以 $\Psi_1\equiv\Psi_2$.

**引理 5.2**　*在引理 5.1 中, 把 (iii) 改成*

(iii)′ *存在 $x_0>0, \pm x_0\in C(\Psi_1)\bigcap C(\Psi_2)$ 使*

$$\int_{[-x_0,x_0)}(1+y^2)d\Psi_1(y)=\int_{[-x_0,x_0)}(1+y^2)d\Psi_2(y),$$

*也有同样的结论:* $\Psi_1\equiv\Psi_2$.

**证** 为证引理 5.2, 只需证明: (i), (ii), (iii)′ ⇒(iii). 事实上, 由 (i) 与 (ii) 可推出 $\Psi_1(A)=\Psi_2(A)$ (当 $0\bar{\in}A, A\in\mathscr{B}^1$). 再用 (iii)′ 知: 对任何 $x>0$, 都有

$$\int_{[-x,x)}(1+y^2)d\Psi_1(y)=\int_{[-x,x)}(1+y^2)d\Psi_2(y),$$

令 $x\downarrow 0$ 即得 $\Psi_1(\{0\})=\Psi_2(\{0\})$.

**引理 5.3** 设 $\Psi_n,\Psi$ 皆为 $\mathscr{B}^1$ 上的有限测度 $(n\geqslant 1)$, 则 $\Psi_n\xrightarrow{\text{w}}\Psi$ 的充要条件是

(i) $\lim\limits_{n\to\infty}\int_{(-\infty,x)}\dfrac{1+y^2}{y^2}d\Psi_n(y)=\int_{(-\infty,x)}\dfrac{1+y^2}{y^2}d\Psi(y)\quad(x<0,x\in C(\Psi))$;

(ii) $\lim\limits_{n\to\infty}\int_{[x,\infty)}\dfrac{1+y^2}{y^2}d\Psi_n(y)=\int_{[x,\infty)}\dfrac{1+y^2}{y^2}d\Psi(y)\quad(x>0,x\in C(\Psi))$;

(iii) $\lim\limits_{n\to\infty}\int_{[-x,x)}(1+y^2)d\Psi_n(y)=\int_{[-x,x)}(1+y^2)d\Psi(y)$
(对一切 $x>0,x\in C(\Psi)$ 或一个这样的 $x$).

**证** 必要性由第三章定理 5.9 可得.

充分性. 因为

$$\begin{aligned}&\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{n\geqslant 1}\int_{|x|>A}d\Psi_n(x)\right)\\ \leqslant&\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{n\geqslant 1}\int_{|x|>A}\frac{1+x^2}{x^2}d\Psi_n(x)\right)\\ \leqslant&\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{1\leqslant n\leqslant N}\int_{|x|>A}\frac{1+x^2}{x^2}d\Psi_n(x)+\sup_{n>N}\int_{|x|>A}\frac{1+x^2}{x^2}d\Psi_n(x)\right),\end{aligned}$$

但是, 由 (i) 和 (ii) 知: 对任何 $\varepsilon>0$, 可选充分大的 $N$ 和 $A_n$,

$$\sup_{n>N}\int_{|x|>A_n}\frac{1+x^2}{x^2}d\Psi_n(x)<\varepsilon,$$

因此

$$\begin{aligned}&\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{n\geqslant 1}\int_{|x|>A}d\Psi_n(x)\right)\\ \leqslant&\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{1\leqslant n\leqslant N}\int_{|x|>A}\frac{1+x^2}{x^2}d\Psi_n(x)+\varepsilon\right)=\varepsilon.\end{aligned}$$

由 $\varepsilon>0$ 可任意小, 得

$$\lim_{\substack{A\to\infty\\ \pm A\in C(\Psi)}}\left(\sup_{n\geqslant 1}\int_{|x|>A}d\Psi_n(x)\right)=0.\tag{5.10}$$

因此, 为证 $\Psi_n \overset{w}{\Longrightarrow} \Psi$, 只需证明 $\Psi_n \overset{lw}{\Longrightarrow} \Psi$. 又因为

$$\begin{aligned}
\Psi_n(\infty) &= \int_{(-\infty,\infty)} d\Psi_n(y) \\
&\leqslant \int_{(-\infty.-x)} \frac{1+y^2}{y^2} d\Psi_n(y) + \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi_n(y) + \int_{[-x,x)} (1+y^2) d\Psi_n(y) \\
&\to \left[\int_{(-\infty,-x)} \frac{1+y^2}{y^2} d\Psi(y) + \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi(y) + \int_{[-x,x)} (1+y^2) d\Psi(y)\right],
\end{aligned}$$

其中 $x > 0, \pm x \in C(\Psi)$, 且使 (i)(ii)(iii) 成立. 所以, $\{\Psi_n(x)\}$ 是一致有界函数列, 从而 $\{\Psi_n\}$ 一定有弱收敛子序列. 因此, 为证 $\{\Psi_n\}$ 局部弱收敛, 只需证明其任一局部弱收敛子序列的极限都是 $\Psi$. 令

$$\Psi_{nk} \overset{lw}{\Longrightarrow} \Phi \quad (k \to \infty),$$

则由 (5.10) 得

$$\Psi_{nk} \overset{w}{\Longrightarrow} \Phi \quad (k \to \infty).$$

所以, 由 [36] 第一章定理 3.10.1 得

$$\begin{aligned}
&\lim_{k\to\infty} \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi_{nk}(y) \\
&= \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Phi(y) \quad (x < 0, x \in C(\Phi)); \\
&\lim_{k\to\infty} \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi_{nk}(y) \\
&= \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Phi(y) \quad (x > 0, x \in C(\Phi)); \\
&\lim_{k\to\infty} \int_{[-x,x)} (1+y^2) d\Psi_{nk}(y) \\
&= \int_{[-x,x)} (1+y^2) d\Phi(y) \quad (x > 0, \pm x \in C(\Phi)).
\end{aligned}$$

把此三式与此引理中三条件 (i)(ii)(iii) 比较, 即可发现 $\Phi$ 与 $\Psi$ 满足引理 5.2 中条件, 故 $\Psi \equiv \Phi$.

**引理 5.4** 把引理 5.3 中的 (iii) 换成

(iii)*
$$\begin{aligned}
&\lim_{x\to 0+} \overline{\lim_{n\to\infty}} \int_{[-x,x)} (1+y^2) d\Psi_n(y) \\
&= \lim_{x\to 0+} \underline{\lim_{n\to\infty}} \int_{[-x,x)} (1+y^2) d\Psi_n(y) = \Psi(\{0\}),
\end{aligned}$$

引理 5.3 的结论仍然成立.

**证** 必要性. 若 $\Psi_n \xrightarrow{\text{w}} \Psi$, 取 $x_m \in C(\Psi), x_m \downarrow 0$, 则

$$\lim_{x_m \downarrow 0} \overline{\lim_{n\to\infty}} \int_{[-x_m,x_m)} (1+y^2)d\Psi_n(y)$$
$$= \lim_{x_m \downarrow 0} \int_{[-x_m,x_m)} (1+y^2)d\Psi(y) = \Psi(\{0\}).$$

而 $\overline{\lim\limits_{n\to\infty}} \int_{[-x,x)}(1+y^2)d\Psi_n(y)$ 是 $x$ 的单调函数, 故

$$\lim_{x\to 0+} \overline{\lim_{n\to\infty}} \int_{[-x,x)} (1+y^2)d\Psi_n(y) = \Psi(\{0\}).$$

仿之可证

$$\lim_{x\to 0+} \underline{\lim_{n\to\infty}} \int_{[-x,x)} (1+y^2)d\Psi_n(y) = \Psi(\{0\}).$$

充分性. 仿引理 5.3, 为证 $\Psi_n \xrightarrow{\text{w}} \Psi$, 只需证明 $\{\Psi_n\}$ 的任何一个局部弱收敛子列的极限都是 $\Psi$. 令 $\{\Psi_{nk}\}$ 是 $\{\Psi_n\}$ 的一个局部弱收敛子列, 其极限为 $\Phi$, 则必有

$$\Psi_{nk} \xrightarrow{\text{w}} \Phi \quad (k\to\infty).$$

因此

$$\lim_{k\to\infty} \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi_{nk}(y)$$
$$= \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Phi(y) \quad (x<0, x\in C(\Phi));$$
$$\lim_{k\to\infty} \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi_{nk}(y)$$
$$= \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Phi(y) \quad (x>0, x\in C(\Phi));$$
$$\lim_{k\to\infty} \int_{[-x,x)} (1+y^2) d\Psi_{nk}(y)$$
$$= \int_{[-x,x)} (1+y^2) d\Phi(y) \quad (x>0, \pm x\in C(\Phi)).$$

把最后一式令 $x \downarrow 0$ 得

$$\lim_{x\downarrow 0} \lim_{k\to\infty} \int_{[-x,x)} (1+y^2)d\Psi_{nk}(y) = \Phi(\{0\}).$$

把此式与 (iii)* 比较, 得 $\Phi(\{0\}) = \Psi(\{0\})$. 总之, 我们证明了对 $\Psi$ 与 $\Phi$ 而言, 满足引理 5.1 中的三个条件, 所以 $\Phi \equiv \Psi$. 引理证毕.

**定理 5.2** 设 $\{f_{nk}(t)\}$ 为 u.a.n. 体系, 则

$$\lim_{n\to\infty} e^{-iA_nt}\prod_k f_{nk}(t) = \{\alpha, \Psi(x)\} \tag{5.11}$$

的充要条件是:

$$\text{(I) (a)}\quad \lim_{n\to\infty}\sum_k F_{nk}(x) = \int_{(-\infty,x)}\frac{1+y^2}{y^2}d\Psi(y)\quad (x<0, x\in C(\Psi)),$$

$$\text{(b)}\quad \lim_{n\to\infty}\sum_k(1-F_{nk}(x)) = \int_{[x,\infty)}\frac{1+y^2}{y^2}d\Psi(y)\quad (x>0, x\in C(\Psi));$$

$$\text{(II)}\quad \lim_{x\to\infty}\sum_k\left\{\int_{|y|<x}y^2dF_{nk}(y) - \left(\int_{|y|<x}ydF_{nk}(y)\right)^2\right\}$$

$$= \int_{|y|<x}(1+y^2)d\Psi(y)\begin{pmatrix}\text{对一切 } x>0, \pm x\in C(\Psi)\\ \text{或一个这样的 } x\end{pmatrix};$$

$$\text{(III)}\quad \lim_{n\to\infty}\left[-A_n + \sum_k a_{nk}(\tau) - \int_{|y|<\tau}yd\Psi(y) + \int_{|y|\geqslant\tau}\frac{1}{y}d\Psi(y)\right]$$

$$= \alpha\begin{pmatrix}\text{对一切}\tau>0, \pm\tau\in C(\Psi),\\ \text{或一个这样的}\tau\end{pmatrix},$$

其中 $F_{nk}(x)$ 是 c.f. $f_{nk}(t)$ 的 d.f., $a_{nk}(\tau)$ 如命题 1.1 后面所定义.

**证** 令

$$\Psi_n(x) = \sum_k\int_{(-\infty,x)}\frac{y^2}{1+y^2}dF_{nk}(y+a_{nk}(\tau)),$$

$$\alpha_n = -A_n + \sum_k a_{nk}(\tau) + \sum_k\int_{\mathbf{R}^1}\frac{x}{1+x^2}dF_{nk}(x+a_{nk}(\tau)),$$

则由定理 5.1 得知 (5.11) 与

$$\begin{cases}\Psi_n(x)\xrightarrow{\text{w}}\Psi(x), & (\text{c}_1)\\ \alpha_n\to\alpha, & (\text{c}_2)\end{cases}$$

等价. 因此, 为证定理 5.2, 只需证明 $(\text{c}_1)$ 及 $(\text{c}_2)$ 与 (I),(II),(III) 等价. 但是, 由引

理 5.3 得知 ($c_1$) 与下列三个条件等价:

$$\sum_k F_{nk}(x+a_{nk}(\tau)) = \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi_n(y)$$

$$\to \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x<0, x\in c(\Psi)), \tag{1°}$$

$$\sum_k (1-F_{nk}(x+a_{nk}(\tau))) = \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi_n(y)$$

$$\to \int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x>0, x\in c(\Psi)), \tag{2°}$$

$$\sum_k \int_{|y|<x} y^2 dF_{nk}(y+a_{nk}(\tau)) = \int_{|y|<x} (1+y^2) d\Psi_n(y)$$

$$\to \int_{|y|<x} (1+y^2) d\Psi(y), \quad \text{对一切 } x>0, \pm x\in c(\Psi), \text{ 或一个这样的 } x. \tag{3°}$$

简记 $a_{nk}(\tau)$ 为 $a_{nk}$. 由于 $\max\limits_k |a_{nk}| \to 0$, 所以对任何正整数 $m$, 都存在 $b_m \in \left(0, \dfrac{1}{m}\right)$, $b_m \downarrow 0, (x\pm b_m)\in C(\Psi)$, 而且存在正整数 $N(m)$, 当 $n \geqslant N(m)$ 时有 $\max\limits_k |a_{nk}| < b_m$. 所以

$$\varlimsup_{n\to\infty} \sum_k F_{nk}(x) \leqslant \varlimsup_{n\to\infty} \sum_k F_{nk}(x+a_{nk}+b_m) \quad (m\geqslant 1),$$

$$\varliminf_{n\to\infty} \sum_k F_{nk}(x) \geqslant \varliminf_{n\to\infty} \sum_k F_{nk}(x+a_{nk}-b_m) \quad (m\geqslant 1).$$

因此, 若 (1°) 成立, 则得

$$\varlimsup_{n\to\infty} \sum_k F_{nk}(x) \leqslant \lim_{b_m\downarrow 0} \int_{(-\infty,x+b_m)} \frac{1+y^2}{y^2} d\Psi(y)$$

$$= \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x<0, x\in c(\Psi)),$$

$$\varliminf_{n\to\infty} \sum_k F_{nk}(x) \geqslant \lim_{b_m\downarrow 0} \int_{(-\infty,x-b_m)} \frac{1+y^2}{y^2} d\Psi(y)$$

$$= \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x<0, x\in c(\Psi)).$$

所以

$$\lim_{n\to\infty} \sum_k F_{nk}(x) = \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x<0, x\in c(\Psi)).$$

此即 (I)(a) 成立. 仿之可证 (2°)⇒(I)(b). 类似地, 可证 (I)(a) 及 (b)⇒(1°) 及 (2°). 总之我们证明了: (I)⇔(1°) 及 (2°), 从而我们证明了

$$(c_1) \Leftrightarrow (\text{I}),\ (3°).$$

所以为证定理, 又只需证

$$(\mathrm{I}),(3^\circ),(\mathrm{c}_2)\Leftrightarrow(\mathrm{I}),(\mathrm{II}),(\mathrm{III}).$$

为此, 我们分几步.

1. $(\mathrm{I})\Rightarrow$“$(3^\circ)\Leftrightarrow(\mathrm{II})$”.

设 (I) 成立. 由不等式 (5.1) 得

$$\begin{aligned}
&\sum_k\int_{|y|<x}y^2dF_{nk}(y+a_{nk})\\
&=\sum_k\left\{\int_{|y|<x}y^2dF_{nk}(y)-\left(\int_{|y|<x}ydF_{nk}(y)\right)^2\right\}+\Delta_n,\\
&|\Delta_n|\leqslant\varepsilon\sum_k\int_{|y|<x}dF_{nk}(y)\\
&+(x+2\varepsilon)^2\sum_k\int_{x-\varepsilon\leqslant|y|<x+\varepsilon}dF_{nk}(y).
\end{aligned}$$

而当 (I) 成立时

$$\begin{aligned}
&\lim_{n\to\infty}\left[\varepsilon\sum_k\int_{|y|<x}dF_{nk}(y)+(x+2\varepsilon)^2\sum_k\int_{x-\varepsilon\leqslant|y|<x+\varepsilon}dF_{nk}(y)\right]\\
&=\varepsilon\int_{|y|<x}\frac{1+y^2}{y^2}d\Psi(y)+(x+2\varepsilon)^2\int_{x-\varepsilon\leqslant|y|<x+\varepsilon}\frac{1+y^2}{y^2}d\Psi(y)
\end{aligned}$$

$(x>0,x,x\pm\varepsilon\in c(\Psi))$. 所以, 由 $\varepsilon>0$ 可以任意小知, 当 (I) 成立时

$$\lim_{n\to\infty}\Delta_n=0.$$

因此, 当 (I) 成立时有

$$(\mathrm{II})\Leftrightarrow(3^\circ).$$

2. $(\mathrm{c}_1)$ (等价地 (I) 及 (II))$\Rightarrow$“$(\mathrm{c}_2)\Leftrightarrow(\mathrm{III})$”. 即在条件 $(\mathrm{c}_1)$ (或 (I) 和 (II), 或 (I) 和 $(3^\circ)$) 下, 有

$$\begin{aligned}
&\sum_k\int_{\mathbf{R}^1}\frac{y}{1+y^2}dF_{nk}(y+a_{nk})\\
&=\int_{|y|\geqslant\tau}\frac{1}{y}d\Psi(y)-\int_{|y|<\tau}yd\Psi(y)+o(1)\quad(n\to\infty).
\end{aligned}\tag{5.12}$$

事实上,

$$
\begin{aligned}
&\sum_k \int_{\mathbf{R}^1} \frac{y}{1+y^2} dF_{nk}(y+a_{nk}) \\
&= \sum_k \int_{|y|<\tau} y dF_{nk}(y+a_{nk}) - \sum_k \int_{|y|<\tau} \frac{y^3}{1+y^2} dF_{nk}(y+a_{nk}) \\
&\quad + \sum_k \int_{|y|\geqslant\tau} \frac{y}{1+y^2} dF_{nk}(y+a_{nk}) \\
&= \sum_k \int_{|y|<\tau} y dF_{nk}(y+a_{nk}) - \int_{|y|<\tau} y d\Psi_n(y) + \int_{|y|\geqslant\tau} \frac{1}{y} d\Psi_n(y),
\end{aligned} \tag{5.13}
$$

利用不等式 (5.6) 及 $(c_1)$ 可得

$$
\begin{aligned}
&\left| \sum_k \int_{|y|<\tau} y dF_{nk}(y+a_{nk}) \right| \\
&\leqslant (\tau+2\varepsilon) \sum_k \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} dF_{nk}(y) + \varepsilon \sum_k \int_{|y|\geqslant\tau} dF_{nk}(y).
\end{aligned}
$$

再用 $(c_1)$ 得

$$
\begin{aligned}
&\lim_{n\to\infty} \left[ (\tau+2\varepsilon) \sum_k \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} dF_{nk}(y) + \varepsilon \sum_k \int_{|y|\geqslant\tau} dF_{nk}(y) \right] \\
&= \lim_{n\to\infty} \left[ (\tau+2\varepsilon) \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} \frac{1+y^2}{y^2} d\Psi_n(y) + \varepsilon \int_{|y|\geqslant\tau} \frac{1+y^2}{y^2} d\Psi_n(y) \right] \\
&= (\tau+2\varepsilon) \int_{\tau-\varepsilon\leqslant|y|<\tau+\varepsilon} \frac{1+y^2}{y^2} d\Psi(y) + \varepsilon \int_{|y|\geqslant\tau} \frac{1+y^2}{y^2} d\Psi(y).
\end{aligned} \tag{5.14}
$$

由 (5.14) 及 $\varepsilon > 0$ 可任意小得知

$$
\sum_k \int_{|y|<\tau} y dF_{nk}(y+a_{nk}) = o(1) \quad (n\to\infty). \tag{5.15}
$$

显然, 由 $(c_1)$ 可知, 当 $n\to\infty$ 时有

$$
\int_{|y|\geqslant\tau} \frac{1}{y} d\Psi_n(y) = \int_{|y|\geqslant\tau} \frac{1}{y} d\Psi(y) + o(1) \quad (\tau>0, \pm\tau\in C(\Psi)), \tag{5.16}
$$

$$
\int_{|y|<\tau} y d\Psi_n(y) = \int_{|y|<\tau} y d\Psi(y) + o(1) \quad (\tau>0, \pm\tau\in C(\Psi)). \tag{5.17}
$$

由 (5.15)、(5.16)、(5.17)、(5.13) 知 (5.12) 成立. 定理证毕.

**定理 5.3** 若 $\{f_{nk}(t)\}$ 是 u.a.n. 体系, 则

$$
\lim_{n\to\infty} e^{-iA_n t} \prod_k f_{nk}(t) = \{\alpha, \Psi(x)\}
$$

的充要条件是定理 5.2 中的 (I),(Ⅲ) 及

$$
\begin{aligned}
(\text{II}_0)\qquad & \lim_{x\to0+}\overline{\lim_{n\to\infty}}\sum_k\left\{\int_{|y|<x}y^2dF_{nk}(y)-\left(\int_{|y|<x}ydF_{nk}(y)\right)^2\right\}\\
&=\lim_{x\to0+}\underline{\lim_{n\to\infty}}\sum_k\left\{\int_{|y|<x}y^2dF_{nk}(y)-\left(\int_{|y|<x}ydF_{nk}(y)\right)^2\right\}\\
&=\Psi(\{0\}).
\end{aligned}
$$

**证**　为证定理 5.3, 只需把定理 5.2 的证明中最初应用引理 5.3 的地方改为应用引理 5.4 即可.

上述两个定理中, $\{A_n\}$ 都是事先给定的实数列, 所以 $\{A_n\}$ 在充要条件中出现.

**定理 5.4**　若 $\{f_{nk}(t)\}$ 为 u.a.n. 体系, 欲有实数列 $\{A_n\}$ 使

$$
\lim_{n\to\infty}e^{-iA_nt}\prod_k f_{nk}(t)=\{\alpha,\Psi(x)\}
$$

的充要条件是 (I),(II) (或者 (I),(II$_0$)). 这时只需取

$$
\begin{aligned}
A_n=&\sum_k a_{nk}(\tau)-\alpha-\int_{|y|<\tau}yd\Psi(y)\\
&+\int_{|y|\geqslant\tau}\frac{1}{y}d\Psi(y)+o(1)\quad(n\to\infty),
\end{aligned}
$$

$(\tau>0,\pm\tau\in C(\Psi))$.

**定理 5.5**　若 $\{f_{nk}(t)\}$ 是 u.a.n. 体系,$f_{nk}(t)$ 是 $X_{nk}$ 的 c.f., 欲有实数列 $\{A_n\}$ 使

$$
\lim_{n\to\infty}e^{-iA_nt}\prod_k f_{nk}(t)=\{\alpha,\Psi(x)\}
$$

的充要条件是定理 5.3 中的 (II) 和

$$
\begin{aligned}
(\text{I}_0)\ (\text{a})\qquad & \lim_{n\to\infty}P\left(\min_k X_{nk}<x\right)\\
&=\begin{cases}1-\exp\left(-\displaystyle\int_{(-\infty,x)}\frac{1+y^2}{y^2}d\Psi(y)\right), & x<0,x\in C(\Psi);\\ 1, & x>0,\end{cases}
\end{aligned}
$$

(b) $$\lim_{n\to\infty} P\left(\max_k X_{nk} < x\right)$$
$$= \begin{cases} 0, & x < 0; \\ \exp\left(-\int_{[x,\infty)} \frac{1+y^2}{y^2} d\Psi(y)\right), & x > 0, x \in C(\Psi). \end{cases}$$

**证** 由定理 5.4, 为证定理 5.5, 只需证明

$$(\text{II}) \Rightarrow \text{“}(\text{I}) \Leftrightarrow (\text{I}_0)\text{”}.$$

只证: (Ⅱ)⇒“(I)(a)⇔($\text{I}_0$)(a)”. 余者类似.

又因为 $x > 0$ 时, 由 u.a.n. 条件总有

$$0 \leqslant 1 - P\left(\min_k X_{nk} < x\right) = P(X_{nk} \geqslant x, k = 1, 2, \cdots, k_n)$$
$$= \prod_k P(X_{nk} \geqslant x) \leqslant \max_k P(|X_{nk}| \geqslant x) \to 0 \quad (n \to \infty),$$

所以

$$\lim_{n\to\infty} P\left(\min_k X_{nk} < x\right) = 1 \quad (x > 0).$$

因此, 欲证在 (Ⅱ) 成立下有 “(I)(a)⇔ ($\text{I}_0$)(a)”, 又只需证在条件 (Ⅱ) 下有

$$\text{“}\lim_{n\to\infty} P\left(\min_k X_{nk} < x\right) = 1 - \exp\left(-\int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y)\right)$$
$$(x < 0, x \in C(\Psi))$$
$$\Longleftrightarrow \lim_{n\to\infty} \sum_k F_{nk}(x) = \int_{(-\infty,x)} \frac{1+y^2}{y^2} d\Psi(y) \quad (x < 0, x \in C(\Psi))\text{”}. \tag{5.18}$$

事实上, 由 u.a.n. 条件有 $\lim\limits_{n\to\infty} \max\limits_k F_{nk}(x) = 0$ (当 $x < 0$), 故

$$P\left(\min_k X_{nk} < x\right) = 1 - \prod_k (1 - F_{nk}(x))$$
$$= 1 - \exp\left(\sum_k \log(1 - F_{nk}(x))\right), \tag{5.19}$$

$$\sum_k F_{nk}(x) \leqslant -\sum_k \log(1 - F_{nk}(x))$$
$$\leqslant \sum_k F_{nk}(x) + \sum_k F_{nk}(x)^2$$
$$\leqslant \sum_k F_{nk}(x)(1 + o(1)) \quad (n \to \infty).$$

因此, 当 $n \to \infty$ 时, $\sum\limits_k F_{nk}(x)$ 与一 $\sum\limits_k \log(1 - F_{nk}(x))$ 有相同的极限 (只要其中有一个极限存在), 因此, 由 (5.19) 得 (5.18), 定理证毕.

## §6 习题及应用

**向 Poisson 特征函数收敛**

1. 试证: 若 $\{f_{nk}(t)\}$ 是 u.a.n. 体系, $\{\alpha, \Psi(x)\} = e^{\lambda(e^{it}-1)}$, 则

$$\lim_{n\to\infty}\prod_k f_{nk}(t) = \{\alpha, \Psi(x)\}$$

的充要条件是:

(I)(a) $\lim\limits_{n\to\infty}\sum\limits_k F_{nk}(x) = 0 \quad (x<0)$,

(b) $\lim\limits_{n\to\infty}\sum\limits_k (1-F_{nk}(x)) = \begin{cases}\lambda, & 0<x<1;\\ 0, & 1<x<\infty;\end{cases}$

(II) $\lim\limits_{n\to\infty}\sum\limits_k \left\{\int_{|y|<x} y^2 dF_{nk}(y) - \left(\int_{|y|<x} y dF_{nk}(y)\right)^2\right\} = 0 \quad (0<x<1)$;

(III) $\lim\limits_{n\to\infty}\sum\limits_k \int_{|y|<\tau} y dF_{nk}(y) = 0 \quad (0<\tau<1)$.

**向正态特征函数收敛**

2. 试证: $\{f_{nk}(t)\}$ 是 u.a.n. 体系且

$$\lim_{n\to\infty} e^{-iA_n t}\prod_k f_{nk}(t) = e^{-\frac{1}{2}t^2}$$

的充要条件是

(I) $\lim\limits_{n\to\infty}\sum\limits_k \int_{|y|\geqslant x} dF_{nk}(y) = 0 \quad (x>0)$;

(II) $\lim\limits_{n\to\infty}\sum\limits_k \left\{\int_{|y|<x} y^2 dF_{nk}(y) - \left(\int_{|y|<x} y dF_{nk}(y)\right)^2\right\} = 1 \quad (x>0)$;

(III) $\lim\limits_{n\to\infty}\left(-A_n + \sum\limits_k \int_{|y|<x} dF_{nk}(y)\right) = 0 \quad (x>0)$.

3. 试证: $\{f_{nk}(t)\}$ 是 u.a.c. 体系且

$$\lim_{n\to\infty} e^{-iA_n t}\prod_k f_{nk}(t) = e^{-\frac{1}{2}t^2}$$

的充要条件是:

(I) $\lim\limits_{n\to\infty}\sum\limits_k \int_{|y|\geqslant x} dF^{\mu}_{nk}(y) = 0 \quad (x>0)$;

(II) $\lim\limits_{n\to\infty}\sum\limits_k \left\{\int_{|y|<x} y^2 dF^{\mu}_{nk}(y) - \left(\int_{|y|<x} y dF^{\mu}_{nk}(y)\right)^2\right\} = 1 \quad (x>0)$;

(III) $\lim\limits_{n\to\infty}\left(-A_n + \sum\limits_k \int_{|y|<\tau} dF^{\mu}_{nk}(y)\right) = 0 \quad (x>0)$.

4. 试证: 设 $\{X_{nk}\}$ 仅满足: 每一行 $X_{n1},\cdots,X_{nk_n}$ 内部是相互独立的随机变量, 且

$$\mathscr{L}\left(\sum_k X_{nk}\right) \xrightarrow{\mathrm{w}} F,$$

$F(x)$ 是 d.f., 则 $\{X_{nk}\}$ 是 u.a.n. 体系且 $F(x)$ 是正态分布的充要条件是

(I)* $$\lim_{n\to\infty}\max_k |X_{nk}|=0,\quad P.$$

**向零一律的特征函数收敛**

5. 试证: $\{f_{nk}(t)\}$ 是 u.a.n. 体系且

$$\lim_{n\to\infty} e^{-iA_nt}\prod_k f_{nk}(t)=1$$

的充要条件是:

(I) $\lim\limits_{n\to\infty}\sum_k \int_{|y|\geqslant x} dF_{nk}(y)=0$ (一切 $x>0$);

(II) $\lim\limits_{n\to\infty}\sum_k \left\{\int_{|y|<x} y^2 dF_{nk}(y)-\left(\int_{|y|<x} ydF_{nk}(y)\right)^2\right\}=0$ (一切或一个 $x>0$);

(III) $\lim\limits_{n\to\infty}\left(-A_n+\sum\limits_k \int_{|y|<\tau} ydF_{nk}(y)\right)=0$, (一切或一个 $\tau>0$).

6. 试证: $\lim\limits_{n\to\infty} e^{-iA_nt}\prod\limits_k f_{nk}(t)=1$ 的充要条件是

$$\begin{cases}\lim\limits_{n\to\infty}\sum\limits_k \boldsymbol{E}\left(\dfrac{(X_{nk}-\mu_{nk})^2}{1+(X_{nk}-\mu_{nk})^2}\right)=0,\\ \lim\limits_{n\to\infty}(-A_n+b_{nk}(\tau))=0.\end{cases}$$

其中 $\mu_{nk}, b_{nk}(\tau)$ 的意义如定理 4.1′.

7. 试证: $\lim\limits_{n\to\infty} e^{-iA_nt}\prod\limits_{k=1}^n f_k\left(\dfrac{t}{B_n}\right)=1$ $(B_n>0)$ 的充要条件是

$$\lim_{n\to\infty}\sum_{k=1}^n \boldsymbol{E}\left(\frac{(X_k-\mu_k)^2}{B_n^2+(X_k-\mu_k)^2}\right)=0,$$

$$\lim_{n\to\infty}\left[-A_n+\frac{1}{B_n}\sum_{k=1}^n\left(\mu_k+\int_{|y|<B_n\tau} ydF_k(y+\mu_k)\right)\right]=0,\quad (\tau>0),$$

其中 $f_k(t), F_k(x), \mu_k$ 分别为 R.V.$X_k$ 的 c.f.,d.f. 和中位数.

8. 试证: $\lim\limits_{n\to\infty} e^{-iA_nt}\left(f\left(\dfrac{t}{B_n}\right)\right)^n=1(B_n>0)$ 的充要条件是

$$\lim_{n\to\infty} n\boldsymbol{E}\left(\frac{(X-\mu)^2}{B_n^2+(X-\mu)^2}\right)=0,$$

$$\lim_{n\to\infty}\left[-A_n+\frac{n}{B_n}\left(\mu+\int_{|y|<B_n\tau} ydF(y+\mu)\right)\right]=0\quad (\tau>0),$$

其中 $f(t), F(x), \mu$ 分别为 R.V.$X$ 的 c.f. 和 d.f. 和中位数.

9. 设 $f(t)$ 是非退化的特征函数, $B_n > 0$ 且

$$\lim_{n\to\infty} e^{-iA_n t}\left(f\left(\frac{t}{B_n}\right)\right)^n = 1,$$

则 $\lim\limits_{n\to\infty}(n/B_n^2) = 0$.

10. 若 $\lim\limits_{n\to\infty}\dfrac{n}{B_n^2} = 0(B_n > 0), f(t)$ 是 c.f., 则

$$\lim_{n\to\infty} e^{-iA_n t}\left(f\left(\frac{t}{B_n}\right)\right)^n = 1$$

的充要条件是

$$\begin{cases} \lim\limits_{n\to\infty} n\displaystyle\int_{\mathbf{R}^1} \frac{x^2}{B_n^2 + x^2} dF(x) = 0, \\ \lim\limits_{n\to\infty}\left(-A_n + \dfrac{n}{B_n}\displaystyle\int_{|x|<B_n\tau} x dF(x)\right) = 0. \end{cases}$$

11. 试证: $\lim\limits_{n\to\infty} e^{-iA_n t}\left(f\left(\dfrac{t}{n^\alpha}\right)\right)^n = 1\left(\alpha > \dfrac{1}{2}\right)$ 的充要条件是

$$\begin{cases} \lim\limits_{z\to\infty} z^{\frac{1}{\alpha}}\displaystyle\int_{|x|\geqslant x} dF(x) = 0, \\ \lim\limits_{n\to\infty}\left[-A_n + n^{1-\alpha}\displaystyle\int_{|x|<n^\alpha} x dF(x)\right] = 0. \end{cases}$$

**稳定特征函数族**

**定义 6.1** 称特征函数 $f(t)$ 是**稳定的**, 如果对任何常数 $a > 0, b > 0$, 都有 $c > 0$ 及 $d$, 使

$$f(at)f(bt) = f(ct)e^{idt}.$$

全部稳定特征函数构成的函数族用 $S$ 表示. 显然退化特征函数和正态特征函数属于 $S$, 且 "$f(t) \in S \Rightarrow e^{iut}f(vt) \in S(\forall u, v \in \mathbf{R}^1)$."

12. 试证下面四族特征函数重合:

$$(1)\ S_1 = \left\{ \tilde{f}(t) \left| \begin{array}{l} \tilde{f}(t) = e^{-iA_n t} f\left(\dfrac{t}{B_n}\right)^n \quad (n \geqslant 1), f \text{ 是 c.f.} \\ \left\{ f_{nk}(t) = f\left(\dfrac{t}{B_n}\right), k = 1, 2, \cdots, n \right\} \text{ 是 u.a.n. 体系.} \end{array} \right. \right\};$$

$$(2)\ S_2 = \left\{ \tilde{f}(t) \left| \begin{array}{l} \tilde{f}(t) = \lim\limits_{n\to\infty} e^{-iA_n t} f\left(\dfrac{t}{B_n}\right)^n \text{ 是 c.f.,} \\ \left\{ f_{nk}(t) = f\left(\dfrac{t}{B_n}\right), k = 1, 2, \cdots, n \right\} \text{ 是 u.a.n. 体系.} \end{array} \right. \right\};$$

$$(3)\ S_3 = \left\{ \tilde{f}(t) \left| \begin{array}{l} \tilde{f}(t) = \lim\limits_{n\to\infty} e^{-iA_n t} f\left(\dfrac{t}{B_n}\right)^n \text{ 是 c.f.,} \\ \left\{ f_{nk}(t) = f\left(\dfrac{t}{B_n}\right), k = 1, 2, \cdots, n \right\} \text{ 是 u.a.c. 体系.} \end{array} \right. \right\};$$

(4) $S_4 = s$.

13. 试求出稳定特征函数 $f(t)$ 的明显的分析表示式, 或者等价地说, 求出它的 Lévy 测度.

# 第十一章　鞅

随机过程论, 已有百年历史. 随机过程论, 按其概率特征来说, 主要有三大类. 一是 Markov 过程论, 其概率特征是: 知道 “现在”、“过去” 和 “将来” 是相互独立的. 二是平稳过程论, 其概率特征是: 其统计规律不依时间的平移而变化. 三是鞅论, 其概率特征是所谓的 “公平博弈” 模型.

什么是 “公平博弈” 模型? 用博弈论中的一个最简单的模型 —— “二人零和” 博弈来说明一下. 假定有甲、乙二人按某种规则 “$A$” 来对弈. 甲、乙二人各有原始资金 $a$ 元和 $b$ 元. 每博弈一局, 若甲胜, 则乙付 1 元给甲, 反之类似. 每博弈一局, 甲胜的概率为 $p$, 乙胜的概率为 $q, p+q=1, p>0, q>0$. 于是每博弈一局, 甲、乙二人获得的总和总是零. 这就是 “二人零和” 博弈的简单通俗的说法. 如果 $p=q=\dfrac{1}{2}$, 博弈一局之后, 甲的资金的数学期望为 $a+(p-q)=a$, 类似地, 乙的资金的数学期望为 $b+(q-p)=b$. 这种博弈规则对甲乙双方来说, 博弈的原则 “$A$” 是公平的. 上面描述的, 就是朴素的 “二人零和” 的 “公平博弈” 模型.

鞅论, 受 “公平博弈” 的启发很大. 鞅的定义, 粗糙地说, 已知时刻 $n$ 以前的资金 $X_0, X_1, \cdots, X_n$ 时, 时刻 $n+1$ 的资金 $X_{n+1}$ 的条件期望就是 $X_n$.

经典鞅论的理论基础, 主要是 J. L. Doob 奠定的, 这是 20 世纪 40 年代的事. 随着调和分析、Banach 空间的几何学和随机分析的发展, 鞅论又得到了长足的发展. D. L. Burkolder, P. A. Meyer 和 J. Neveu 等人都做出过巨大的贡献.

由于本书是讲高等概率论及其应用, 而非专讲随机过程论, 所以平稳过程被略去了, 而只讲了随机过程论中用得较多的可数状态的 Markov 过程和经典鞅

论, 而这两方面文献亦浩如烟海, 我们只摘要地讲一些最基本的.

## §1 鞅的基本概念及其不等式

前面我们已经介绍了朴素的鞅的概念及其粗略历史, 提到了鞅的概念与公平博弈的关系, 介绍了经典鞅论与近代鞅论的一些代表性人物.

在 §1 中, 除了给出鞅的精确定义以外, 还举了一些简单例子. 特别是介绍了一些重要的鞅不等式, 如 Doob 不等式、Kolmogorov 不等式等.

**定义 1.1** 设 $(\Omega,\mathscr{F},P)$ 为概率空间, $\boldsymbol{T}\subset\overline{\mathbf{R}},\{\mathscr{F}_1:t\in\boldsymbol{T}\}$ 是 $\mathscr{F}$ 中的一族单调非降子 $\sigma$ 代数,$\{X_t:t\in\boldsymbol{T}\}$ 是一个适应于 $\{\mathscr{F}_t:t\in\boldsymbol{T}\}$ 的以 $(\mathbf{R},\mathscr{B}^1)$ 为状态空间的随机过程, $\boldsymbol{E}(|X_t|)<\infty(t\in\boldsymbol{T})$ (此处及今后,$\mathscr{B}^d\overset{\text{def.}}{=\!=}\mathscr{B}(\mathbf{R}^d)$), 如果

$$\boldsymbol{E}(X_t|\mathscr{F}_s)\leqslant X_s\quad(s\leqslant t,s,t\in\boldsymbol{T}),\tag{1.1}$$

则称 $X=\{X_t,\mathscr{F}_t,t\in\boldsymbol{T}\}$ 是**上鞅**. 若 (1.1) 式代之以

$$\boldsymbol{E}(X_t|\mathscr{F}_s)\geqslant X_s\quad(s\leqslant t,s,t\in\boldsymbol{T}),\tag{1.2}$$

则称 $X$ 为**下鞅**. 既是上鞅又是下鞅者, 称之为**鞅**.. 本章总设概率空间是完备的.

**例 1.1** 设 $(\Omega,\mathscr{F},P),\boldsymbol{T},\{\mathscr{F}_t:t\in\boldsymbol{T}\}$ 如定义 1.1, $\xi$ 是一随机变量, $\boldsymbol{E}(|\xi|)<\infty$, 则 $X=\{X_t=\boldsymbol{E}(\xi|\mathscr{F}_t),\mathscr{F}_t,t\in\boldsymbol{T}\}$ 是一个鞅.

**例 1.2** 设 $(\Omega,\mathscr{F},P)$ 是一概率空间, $\boldsymbol{T}=\{0,1,\cdots\},\{X_n:n\in\boldsymbol{T}\}$ 是 $(\Omega,\mathscr{F},P)$ 上的一个自由随机徘徊, 即

$$X_n=\sum_{k=0}^{n}\xi_k(n\in\boldsymbol{T}),\quad\{\xi_k:k\in\boldsymbol{T}\}$$

是相互独立具有共同分布函数的随机变量序列:

$$P(\xi_k=1)=p,\quad P(\xi_k=-1)=q(k\in\boldsymbol{T}),$$
$$0<p<1,\quad p+q=1.$$

令 $\mathscr{F}_n=\sigma(X_s,0\leqslant s\leqslant n),X=(X_n,\mathscr{F}_n,n\in\boldsymbol{T})$, 则

(1) $X$ 是鞅 $\Leftrightarrow p=q$;

(2) $X$ 是下鞅 $\Leftrightarrow p\geqslant q$;

(3) $X$ 是上鞅 $\Leftrightarrow p\leqslant q$.

事实上, “$\boldsymbol{E}(X_{n+1}|\mathscr{F}_n)=X_n$ (对应地 $\geqslant$ 或 $\leqslant$)$\Leftrightarrow X$ 是鞅 (对应地, 下鞅或上鞅)”. 而

$$\begin{aligned}\boldsymbol{E}(X_{n+1}|\mathscr{F}_n)&=\boldsymbol{E}(X_n|\mathscr{F}_n)+\boldsymbol{E}(\xi_{n+1}|\mathscr{F}_n)\\&=X_n+\boldsymbol{E}(\xi_{n+1})=X_n+(p-q)\quad(n\geqslant 0),\end{aligned}$$

故论断成立.

**例 1.3** 令 $\Omega=[0,1)$, $\mathscr{F}$ 是 $\Omega$ 中全体 Borel 子集构成的 $\sigma$ 代数, $P$ 是 $\mathscr{F}$ 上的 Lebesgue 测度, 遂得概率空间 $(\Omega,\mathscr{F},P)$. 令 $\xi$ 是 $(\Omega,\mathscr{F},P)$ 上的实值随机变量, 且 $\boldsymbol{E}(|\xi|)<\infty$. 把 $\xi$ 的定义域由 $[0,1)$ 按下法扩张到 $[0,2)$ 上,

$$\xi(x+1)=\xi(x),\quad x\in[0,1).$$

再令

$$\eta_n(x)=\frac{1}{n}\sum_{k=0}^{n-1}\xi\left(x+\frac{k}{n}\right),\quad x\in[0,1),$$

$\mathscr{F}_n'=\{A:A\in\mathscr{F},A$ 是以 $\frac{1}{n}$ 为周期的集合 $\}$, $\mathscr{F}_{1/n}=\mathscr{F}_{2^n}'$. 所谓 $A$ 是以 $\frac{1}{n}$ 为周期的集合, 即 "$x\in A\Leftrightarrow\left(x+\frac{1}{n}-\left[x+\frac{1}{n}\right]\right)\in A$", $[\alpha]$ 表示不大于 $\alpha$ 的最大整数. 则 $X=\{X_{1/n}=\boldsymbol{E}(\xi|\mathscr{F}_{1/n}),\mathscr{F}_{1/n},n\geqslant1\}$ 是鞅, 且 $X_{1/n}=\boldsymbol{E}(\xi|\mathscr{F}_{1/n})=\eta_{2^n}$.

首先证明 $\mathscr{F}_n$ 是 $\sigma$ 代数, 显然 $\Omega\in\mathscr{F}_n'$. 又若 $A_i\in\mathscr{F}_n'(i=1,2,\cdots)$, 则 $\bigcup_{i=1}^{\infty}A_i\in\mathscr{F}$, 且

$$x\in\bigcup_{i=1}^{\infty}A_i\Leftrightarrow \text{存在 } i_0\geqslant1, \text{使 }\left(x+\frac{1}{n}\right)-\left[x+\frac{1}{n}\right]\in A_{i_0}$$
$$\Leftrightarrow\left(x+\frac{1}{n}\right)-\left[x+\frac{1}{n}\right]\in\bigcup_{i=1}^{\infty}A_i,$$

即 $\bigcup_{i=1}^{\infty}A_i\in\mathscr{F}_n'$. 仿之可证 $A,B\in\mathscr{F}_n'\Rightarrow A-B\in\mathscr{F}_n'$. 所以 $\mathscr{F}_n'$ 是 $\sigma$ 代数.

显然 $\{\mathscr{F}_{1/n}=\mathscr{F}_{2^n}':n\geqslant1\}$ 是单调非降的. 所以, 由例 1.1 得知 $X$ 是鞅.

其次证明: $\eta_n=\boldsymbol{E}(\xi|\mathscr{F}_n')$. 事实上, 由 $\xi(x)=\xi(x+1)(x\in[0,1))$, 可得

$$\begin{aligned}\eta_n\left(x+\frac{1}{n}-\left[x+\frac{1}{n}\right]\right)&=\frac{1}{n}\sum_{k=0}^{n-1}\xi\left(x+\frac{k+1}{n}-\left[x+\frac{1}{n}\right]\right)\\&=\begin{cases}\dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1}\xi\left(x+\frac{k+1}{n}-1\right), & \text{当 } x\in\left[\dfrac{n-1}{n},\dfrac{n}{n}\right),\\ \dfrac{1}{n}\displaystyle\sum_{k=0}^{n-1}\xi\left(x+\frac{k+1}{n}\right), & \text{当 } x\in\left[0,\dfrac{n-1}{n}\right),\end{cases}\\&=\frac{1}{n}\sum_{k=0}^{n-1}\xi\left(x+\frac{k+1}{n}\right)=\frac{1}{n}\sum_{k=0}^{n-1}\xi\left(x+\frac{k}{n}\right)=\eta_n(x).\end{aligned}$$

若令

$$A=\{x:x\in\Omega,\eta_n(x)<\lambda\},\quad \lambda\text{为实数},$$

则有

$$x \in A \Leftrightarrow \eta_n(x) < \lambda \Leftrightarrow \eta_n\left(x + \frac{1}{n} - \left[x + \frac{1}{n}\right]\right) < \lambda$$
$$\Leftrightarrow x + \frac{1}{n} - \left[x + \frac{1}{n}\right] \in A.$$

显然 $A \in \mathscr{F}$, 所以 $A \in \mathscr{F}'_n$, 即 $\eta_n \in \mathscr{F}'_n$. 再证对任何 $B \in \mathscr{F}'_n$, 有

$$\int_B \eta_n dP = \int_B \xi dP. \tag{1.3}$$

令

$$B\left(-\frac{k}{n}\right) = \left\{x : x = y - \frac{k}{n}, y \in B, y - \frac{k}{n} \geqslant 0\right\}$$
$$\bigcup \left\{x : x = y - \frac{k}{n} + 1, y \in B, y - \frac{k}{n} < 0\right\}, \tag{1.4}$$

若注意

$$x = \begin{cases} \left(\left(x + \dfrac{k}{n}\right) - \left[x + \dfrac{k}{n}\right]\right) - \dfrac{k}{n}, \\ \qquad \text{当 } x \in [0,1), x + \dfrac{k}{n} < 1; \\ \left(\left(x + \dfrac{k}{n}\right) - \left[x + \dfrac{k}{n}\right]\right) - \dfrac{k}{n} + 1, \\ \qquad \text{当 } x \in [0,1), x + \dfrac{k}{n} \geqslant 1, \end{cases}$$

$x \in [0,1), x + \dfrac{k}{n} < 1 \Leftrightarrow x \in [0,1), x - \left[x + \dfrac{k}{n}\right] \geqslant 0$, 则有

$$x = \begin{cases} \left(\left(x + \dfrac{k}{n}\right) - \left[x + \dfrac{k}{n}\right]\right) - \dfrac{k}{n}, \\ \qquad \text{当 } x \in [0,1), x - \left[x + \dfrac{k}{n}\right] \geqslant 0; \\ \left(\left(x + \dfrac{k}{n}\right) - \left[x + \dfrac{k}{n}\right]\right) - \dfrac{k}{n} + 1, \\ \qquad \text{当 } x \in [0,1), x - \left[x + \dfrac{k}{n}\right] < 0. \end{cases}$$

由 $B \in \mathscr{F}'_n$ 及 $B\left(-\dfrac{k}{n}\right)$, $\mathscr{F}'_n$ 的定义可得

$$x \in B\left(-\frac{k}{n}\right) \Leftrightarrow \left(x + \frac{k}{n}\right) - \left[x + \frac{k}{n}\right] \in B \Leftrightarrow x \in B,$$

即

$$B\left(-\frac{k}{n}\right) \equiv B, \quad k = 0, 1, \cdots, n-1. \tag{1.5}$$

所以由 $\eta_n$ 及 $B\left(-\frac{k}{n}\right)$ 的定义和 (1.5) 式得

$$\begin{aligned}\int_B \eta_n dP &= \frac{1}{n}\sum_{k=0}^{n-1}\int_B \xi\left(x+\frac{k}{n}\right)dP\\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int_{B(-\frac{k}{n})}\xi(x)dP\\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int_B \xi(x)dP = \int_B \xi dP.\end{aligned}$$

这就证明了 $\eta_n = \boldsymbol{E}(\xi|\mathscr{F}'_n)$. 且有

$$X_{1/n} = \boldsymbol{E}(\xi|\mathscr{F}_{1/n}) = \boldsymbol{E}(\xi|\mathscr{F}'_{2^n}) = \eta_{2^n}.$$

**例 1.4** 设 $\{X_t : t \in [0,\infty)\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的 S.B$^1$.M.O., 令 $\mathscr{G}_t^0 = \sigma(X_s, s \leqslant t), t \in [0,\infty)$, 则

(1) $\{X_t, \mathscr{G}_t^0, t \in [0,\infty)\}$ 是鞅;

(2) $\{X_t^2 - t, \mathscr{G}_t^0, t \in [0,\infty)\}$ 是鞅;

(3) $\{e^{aX_t - \frac{a^2 t}{2}}, \mathscr{G}_t^0, t \in [0,\infty)\}$ 是鞅.

**证** 由于 $\{X_t : t \in [0,\infty)\}$ 具有独立增量, 若注意

$$\sigma(X_{t_1}, X_{t_2}, \cdots, X_{t_n}) = \sigma(X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}})$$

对一切 $0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n$ 成立, 则可知

$$\text{“}t \geqslant s \Rightarrow \sigma(X_t - X_s) \text{ 与 } \mathscr{G}_s^0 \text{ 独立”.}$$

所以

(1) $\boldsymbol{E}(X_t|\mathscr{G}_s^0) = \boldsymbol{E}(X_s|\mathscr{G}_s^0) + \boldsymbol{E}(X_t - X_s|\mathscr{G}_s^0)$

$$= X_s + \boldsymbol{E}(X_t - X_s) = X_s \quad (s \leqslant t),$$

故 $\{X_t, \mathscr{G}_t^0, t \in [0,\infty)\}$ 是鞅.

$$(2)\ \boldsymbol{E}(X_t^2-t|\mathscr{G}_s^0)=\boldsymbol{E}((X_s+X_t-X_s)^2|\mathscr{G}_s^0)-t$$

$$\begin{aligned}&=\boldsymbol{E}(X_s^2|\mathscr{G}_s^0)+2X_s\boldsymbol{E}(X_t-X_s|\mathscr{G}_s^0)\\&\quad+\boldsymbol{E}((X_t-X_s)^2|\mathscr{G}_s^0)-t\\&=X_s^2-2X_s\boldsymbol{E}(X_t-X_s)+\boldsymbol{E}((X_t-X_s)^2)-t\\&=X_s^2-0+(t-s)-t\\&=X_s^2-s\quad(s\leqslant t),\end{aligned}$$

所以 $\{X_t^2-t,\mathscr{G}_t^0,t\in[0,\infty)\}$ 是鞅.

(3) 若 $Z\sim N(0,\sigma^2)$, 即 $Z$ 服从期望为 0、方差为 $\sigma^2$ 的正态分布, 则 $\boldsymbol{E}(e^{i\alpha Z})=e^{-\frac{\sigma^2\alpha^2}{2}},\boldsymbol{E}(e^{\alpha Z})=e^{\frac{\sigma^2\alpha^2}{2}}$, 所以由 $(X_t-X_s)\sim N(0,t-s)$ 得

$$\begin{aligned}\boldsymbol{E}(e^{aX_t-\frac{a^2t}{2}}|\mathscr{G}_s^0)&=\boldsymbol{E}(e^{a(X_t-X_s)}e^{aX_s-\frac{a^2t}{2}}|\mathscr{G}_s^0)\\&=e^{aX_s-\frac{a^2t}{2}}\boldsymbol{E}(e^{a(X_t-X_s)}|\mathscr{G}_s^0)\\&=e^{aX_s-\frac{a^2t}{2}}\boldsymbol{E}(e^{a(X_t-X_s)})\\&=e^{aX_s-\frac{a^2t}{2}}e^{\frac{a^2(t-s)}{2}}=e^{aX_s-\frac{a^2s}{2}}\quad(s\leqslant t).\end{aligned}$$

所以 $\{e^{X_t-\frac{a^2t}{2}},\mathscr{G}_t^0,t\in[0,\infty)\}$ 是鞅.

对于 (上、下) 鞅 $X=\{X_t,\mathscr{F}_t,t\in\boldsymbol{T}\}$ 的 "时间" 参数集 $\boldsymbol{T}$, 我们最感兴趣的是两种特例: (1) $\boldsymbol{T}=\{0,1,2,\cdots\}$; (2) $\boldsymbol{T}=[0,\infty)$. 对于特例 (1). 称 $X$ 为离散时间 (上、下) 鞅, 对于特例 (2), 称 $X$ 为连续时间 (上、下) 鞅.

**命题 1.1** 设 $\{X_n:n=0,1,\cdots\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的 $\{\mathscr{F}_n:n=0,1,\cdots\}$ 适应随机过程, $\boldsymbol{E}(|X_n|)<\infty,n=0,1,\cdots$, 则 $X=\{X_n,\mathscr{F}_n,n=0,1,\cdots\}$ 是鞅的充要条件是:

$$\boldsymbol{E}(X_{n+1}|\mathscr{F}_n)=X_n\quad(n\geqslant 0).\tag{1.6}$$

**证** 必要性显然成立, 现证充分性. 设 (1.6) 式成立, 则对任意 $m>0,n\geqslant 0$, 有

$$\begin{aligned}\boldsymbol{E}(X_{n+m}|\mathscr{F}_n)&=\boldsymbol{E}(\boldsymbol{E}(X_{n+m}|\mathscr{F}_{n+m-1})|\mathscr{F}_n)\\&=\boldsymbol{E}(X_{n+m-1}|\mathscr{F}_n)=\cdots\\&=\boldsymbol{E}(X_{n+1}|\mathscr{F}_n)=X_n.\end{aligned}$$

所以 $X=\{X_m,\mathscr{F}_n,n=0,1,\cdots\}$ 是鞅.

**定义 1.2**　设 $(\Omega,\mathscr{F},\mu)$ 为任一测度空间,$1\leqslant p<\infty$, 记

$$L^p(\Omega,\mathscr{F},\mu)=\left\{f:f\in\mathscr{F},\int_\Omega |f|^p d\mu<\infty\right\},\tag{1.7}$$

$L^\infty(\Omega,\mathscr{F},\mu)=\{f:f\in\mathscr{F}$, 且存在实数 $\alpha$, 使 $|f|\leqslant\alpha,\text{a.s.}\}$. 若 $f,g\in L^2(\Omega,\mathscr{F},\mu)$, $\int_\Omega fgd\mu=0$, 则称 $f$ 与 $g$ **互相垂直**, 记作 $f\perp g$; 若 $f_n\in L^2(\Omega,\mathscr{F},\mu)$ $(n\geqslant 0)$, 且 $f_n\perp f_m(n\neq m)$, 则称 $\{f_n:n\geqslant 0\}$ 是正交系.

对 $f\in L^p(\Omega,\mathscr{F},\mu)$, 记

$$\|f\|_p=\left(\int_\Omega |f|^p d\mu\right)^{-\frac{1}{p}};$$

对 $f\in L^\infty(\Omega,\mathscr{F},\mu)$, 记 $\|f\|_\infty=\inf\{\alpha:|f|\leqslant\alpha,\ \text{a.s.}\}$

**定义 1.3**　设 $(\Omega,\mathscr{F},\mu)$ 是任一测度空间, $\Gamma$ 是任一指标集合, $1\leqslant p\leqslant\infty$, 对每一个 $t\in\Gamma$, 有

$$f_t:\Omega\mapsto\overline{\mathbf{R}},\quad f_t\in\mathscr{F},$$

称 $f=\{f_t:t\in\Gamma\}$ 在 $(\Omega,\mathscr{F},\mu)$ 上是 **$L^p$ 有界的** (简称 $L^p$ 有界的) 或 $f\in L^p(\Omega,\mathscr{F},\mu)$, 如果

$$\sup_{t\in\Gamma}\|f_t\|_p<\infty,\tag{1.8}$$

记 $\|f\|_p=\sup\limits_{t\in\Gamma}\|f_t\|_p$.

**定义 1.4**　设 $f=\{f_t:t\in\Gamma\}$ 是测度空间 $(\Omega,\mathscr{F},\mu)$ 上一族 $\mathscr{F}$ 可测函数, 如果

$$\lim_{k\to\infty}\sup_{t\in\Gamma}\int_{\{|f_t|\geqslant k\}}|f_t|du=0,\tag{1.9}$$

则称 $\{f_t:t\in\Gamma\}$ 在 $(\Omega,\mathscr{F},\mu)$ 上是**一致可积的**, 简称一致可积.

下面我们先研究离散时间鞅.

首先研究离散时间鞅 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 与 "**鞅差序列**" $\{D_n=X_n-X_{n-1},X_{-1}\equiv 0,n\geqslant 0\}$ 的关系.

**定理 1.1**　设 $(\Omega,\mathscr{F},P)$ 为概率空间, $D_n\in L^1(\Omega,\mathscr{F},P),n=0,1,\cdots$.

(1) 若 $X_n=\sum\limits_{k=0}^{n}D_k,\{\mathscr{F}_n:n\geqslant 0\}$ 是 $\mathscr{F}$ 中一族非降子 $\sigma$ 代数族, $D_n\in\mathscr{F}_n$, 则

$$X=\{X_n,\mathscr{F}_n,n\geqslant 0\}\text{ 是鞅 }\Leftrightarrow \boldsymbol{E}(D_n|\mathscr{F}_{n-1})=0\quad(n\geqslant 1).$$

(2) 若令 $\mathscr{F}_n=\sigma(D_0,D_1,\cdots,D_n)(n\geqslant 0)$, 则 $\{D_n:n\geqslant 0\}$ 是某个鞅差序列, 即存在鞅 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$, 使 $D_n=X_n-X_{n-1}(n\geqslant 0,X_{-1}\overset{\text{def.}}{=\!=}0)$ 的充

要条件是:

$$\varphi(D_0, D_1, \cdots, D_{n-1}) \perp D_n(n \geqslant 1, \varphi \in b\mathscr{B}^n,$$
$$b\mathscr{B}^n \text{ 是有界的 } n \text{ 维 Borel 可测函数全体}). \tag{1.10}$$

(3) 若 $X_n = \sum_{k=0}^{n} D_k, \mathscr{F}_n = \sigma(D_0, D_1, \cdots, D_n)(n \geqslant 0), X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^2$ 有界鞅, 则 $\{D_n : n \geqslant 0\}$ 是 $(\Omega, \mathscr{F}, P)$ 上正交系, 且

$$\varphi(D_0, D_1, \cdots, D_{n-1}) \perp D_n \quad (n \geqslant 1, \varphi \in b\mathscr{B}^n).$$

**证** (1) 由

$$\begin{aligned} \boldsymbol{E}(X_n|\mathscr{F}_{n-1}) &= \boldsymbol{E}(D_n|\mathscr{F}_{n-1}) + \boldsymbol{E}(X_{n-1}|\mathscr{F}_{n-1}) \\ &= \boldsymbol{E}(D_n|\mathscr{F}_{n-1}) + X_{n-1} \end{aligned}$$

即得.

(2) 必要性. 若存在鞅 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$, 使 $X_n - X_{n-1} = D_n(n \geqslant 0, X_{-1} \equiv 0)$, 则由 (1) 得 $\boldsymbol{E}(D_n|\mathscr{F}_{n-1}) = 0(n \geqslant 1)$. 因此, 对任何 $\varphi \in b\mathscr{B}^n$, 有

$$\begin{aligned} 0 &= \boldsymbol{E}(\varphi(D_0, D_1, \cdots, D_{n-1})\boldsymbol{E}(D_n|\mathscr{F}_{n-1})) \\ &= \boldsymbol{E}(\boldsymbol{E}(\varphi(D_0, D_1, \cdots, D_{n-1})D_n|\mathscr{F}_{n-1})) \\ &= \boldsymbol{E}(\varphi(D_0, D_1, \cdots, D_{n-1})D_n) \quad (n \geqslant 1), \end{aligned}$$

即 (1.10) 式成立.

充分性. 设 (1.10) 式成立. 任取 $B \in \mathscr{F}_{n-1}, h = \mathbf{1}_B$, 则由 Doob 复合函数定理, 必存在 $\varphi \in b\mathscr{B}^n$, 使 $h = \varphi(D_0, D_1, \cdots, D_{n-1})$, 所以由 (1.10) 式得

$$\boldsymbol{E}(hD_n) = 0,$$

亦即 $\int_B D_n dP = 0(B \in \mathscr{F}_{n-1}, n \geqslant 1)$. 所以

$$\boldsymbol{E}(D_n|\mathscr{F}_{n-1}) = 0 \quad (n \geqslant 1).$$

因此, 由 (1) 得知 $X = \{X_n, \mathscr{F}_n, n \geqslant 1\}$ 是鞅.

(3) 由 (2) 得知

$$\varphi(D_0, D_1, \cdots, D_{n-1}) \perp D_n \quad (\varphi \in b\mathscr{B}^n, n \geqslant 1).$$

又因为 $X$ 是 $L^2$ 有界的, 所以 $D_n = (X_n - X_{n-1}) \in L^2(\Omega, \mathscr{F}, P)$, 且对任何 $n < m$, 有

$$\boldsymbol{E}(D_n D_m) = \boldsymbol{E}(D_m \boldsymbol{E}(D_n|\mathscr{F}_m)) = 0.$$

故 (3) 得证.

**命题 1.2** 设 $\{X_n,\mathscr{F}_n,n=0,1,\cdots\},\{Y_n,\mathscr{F}_n,n=0,1,\cdots\}$ 为鞅 (相应地, 下鞅), 则 $\{X_n+Y_n,\mathscr{F}_n,n=0,1,\cdots\}$ 为鞅 (相应地, 下鞅), $\{X_n\wedge Y_n,\mathscr{F}_n,n=0,1,\cdots\}$ 为上鞅.

**证** 显然成立.

**命题 1.3** 设 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为鞅 (相应地, 下鞅), $f$ 为 $\mathbf{R}$ 上的凸函数 (相应地, 非降凸函数), $\boldsymbol{E}(|f(X_n)|)<\infty(n\geqslant 0)$, 则 $\{f(X_n),\mathscr{F}_n,n\geqslant 0\}$ 为下鞅.

**证** 若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为鞅, $f$ 为 $\mathbf{R}$ 上的凸函数, 则用 Jensen 不等式得

$$\boldsymbol{E}(f(X_{n+1})|\mathscr{F}_n)\geqslant f(\boldsymbol{E}(X_{n+1}|\mathscr{F}_n))=f(X_n)\quad(n\geqslant 0),$$

所以 $\{f(X_n),\mathscr{F}_n,n\geqslant 0\}$ 是下鞅.

若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是下鞅, $f$ 是 $\mathbf{R}$ 上的非降凸函数, 由 Jensen 不等式及 $f$ 的非降性得

$$\boldsymbol{E}(f(X_{n+1})|\mathscr{F}_n)\geqslant f(\boldsymbol{E}(X_{n+1}|\mathscr{F}_n))\geqslant f(X_n)\quad(n\geqslant 0),$$

所以 $\{f(X_n),\mathscr{F}_n,n\geqslant 0\}$ 是下鞅.

**命题 1.4** (1) 若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为鞅 (或非负下鞅),

$$\boldsymbol{E}(|X_n|^p)<\infty\quad(n\geqslant 0),$$

则 $\{|X_n|^p,\mathscr{F}_n,n\geqslant 0\}$ 为下鞅 $(1\leqslant p<\infty)$.

(2) 若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为下鞅, 则 $\{X_n^+,\mathscr{F}_n,n\geqslant 0\}$ 亦然 (其中 $X_n^+=X_n\vee 0$).

**证** (1) 若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为鞅, 由 $f(x)=|x|^p(1\leqslant p<\infty)$ 是 $\mathbf{R}$ 上的凸函数, 由命题 1.3 得知 $\{|X_n|^p,\mathscr{F}_n,n\geqslant 0\}$ 是下鞅.

若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为非负下鞅, 由于 $f(x)=x^p(1\leqslant p<\infty)$ 是 $[0,\infty)$ 上的非降凸函数, 仍用命题 1.3 得知 $\{X_n^p,\mathscr{F}_n,n\geqslant 0\}$ 为下鞅.

(2) 若 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为下鞅, 令 $f(x)=x^+$, 则 $f$ 是 $\mathbf{R}$ 上的非降凸函数, 用命题 1.3 得 $\{X_n^+=f(X_n),\mathscr{F}_n,n\geqslant 0\}$ 为下鞅.

**定理 1.2** 设 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的上鞅, 则对任何

$\lambda > 0, n \geqslant 0$, 有

$$\lambda P\left(\sup_{k\leqslant n} X_k > \lambda\right) \leqslant \boldsymbol{E}(X_0) - \int_{\left\{\sup\limits_{k\leqslant n} X_k < \lambda\right\}} X_n dP; \tag{1.11}$$

$$\lambda P\left(\inf_{k\leqslant n} X_k < -\lambda\right) \leqslant \int_{\left\{\inf\limits_{k\leqslant n} X_k < -\lambda\right\}} (-X_n) dP; \tag{1.12}$$

$$\lambda P\left(\sup_{k\leqslant n} |X_k| > \lambda\right) \leqslant \boldsymbol{E}(X_0) + 2\boldsymbol{E}(X_n^-), \tag{1.13}$$

其中 $X_n^- = (-X_n) \vee 0$.

若 $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅或非负下鞅, 则

$$\lambda P(X_n^* > \lambda) \leqslant \int_{|X_n^* > \lambda|} |X_n| dP; \tag{1.13$'$}$$

$$\lambda P(X^* > \lambda) \leqslant \sup_{n\geqslant 0} \boldsymbol{E}(|X_n|), \tag{1.13$''$}$$

其中, $X_n^* = \sup\limits_{k\leqslant n} |X_k|; X^* = \sup\limits_{n\geqslant 0} X_n^*$.

**证** 令 $A_0 = \{X_0 < -\lambda\}, A_k = \{X_0 \geqslant -\lambda, X_1 \geqslant -\lambda, \cdots, X_{k-1} \geqslant -\lambda, X_k < -\lambda\}(k \geqslant 1)$, 则 $A_k \in \mathscr{F}_k(k \geqslant 0), \{A_k\}$ 不交. 而由 $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是上鞅得 $\boldsymbol{E}(X_n|\mathscr{F}_k) \leqslant X_k(n \geqslant k \geqslant 0)$, 所以

$$\int_{A_k} X_n dP \leqslant \int_{A_k} X_k dP \quad (n \geqslant k \geqslant 0).$$

因此,

$$\begin{aligned}
\lambda P\left(\inf_{k\leqslant n} X_k < -\lambda\right) &= \lambda P\left(\bigcup_{k=0}^n A_k\right) = \lambda \sum_{k=0}^n P(A_k) \\
&= \sum_{k=0}^n \int_{A_k} \lambda dP \leqslant \sum_{k=0}^n \int_{A_k} -X_k dP \\
&\leqslant \sum_{k=0}^n \int_{A_k} -X_n dP = -\int_{\bigcup\limits_{k=0}^n A_k} X_n dP \\
&= \int_{\{\inf\limits_{k\leqslant n} X_k < -\lambda\}} -X_n dP.
\end{aligned}$$

故 (1.12) 式得证.

下证 (1.11) 式. 令 $\tau = \inf\{k : k \geqslant 0, X_k > \lambda\} \wedge n, \tau_k = \tau \wedge k (k = 0, 1, \cdots, n)$, 则 $0 = \tau_0 \leqslant \tau_1 \leqslant \cdots \leqslant \tau_n = \tau, 0 \leqslant \tau_k - \tau_{k-1} \leqslant 1, \tau_k$ 皆为 $\{\mathscr{F}_n\}$ 停时, 所以由本

章命题 3.4 可推出

$$\mathscr{F}_0 = \mathscr{F}_{\tau_1} \subset \mathscr{F}_{\tau_1} \subset \cdots \subset \mathscr{F}_{\tau_n} = \mathscr{F}_{\tau},$$

且 $X_{\tau_k} \in \mathscr{F}_{\tau_k}$. 任取 $A \in \mathscr{F}_{\tau_k}(0 \leqslant k < n)$, 必有

$$A \bigcap \{\tau_k = j\} \bigcap \{\tau_{k+1} > j\} \in \mathscr{F}_j,$$

所以由上鞅性得

$$\int_A (X_{\tau_k} - X_{\tau_{k+1}})dP = \sum_{j=0}^{k} \int_{A\cap\{\tau_k=j\}\cap\{\tau_{k+1}>j\}} (X_j - X_{j+1})dP \geqslant 0.$$

所以, 若 $A \in \mathscr{F}_0 \subset \mathscr{F}_{\tau_k}(k = 0, 1, \cdots, n)$, 则

$$\int_A (X_0 - X_\tau)dP = \sum_{k=0}^{n-1} \int_A (X_{\tau_k} - X_{\tau_{k+1}})dP \geqslant 0,$$

而 $X_0 \in \mathscr{F}_0$, 所以由

$$X_0 \geqslant E(X_\tau|\mathscr{F}_0), \tag{1.11}''$$

即得

$$\begin{aligned}\boldsymbol{E}(X_0) \geqslant \boldsymbol{E}(X_\tau) &= \int_{\left\{\sup\limits_{k\leqslant n} X_k > \lambda\right\}} X_\tau dP + \int_{\left\{\sup\limits_{k\leqslant n} X_k \leqslant \lambda\right\}} X_\tau dP \\ &\geqslant \lambda P\left(\sup_{k\leqslant n} X_k > \lambda\right) + \int_{\left\{\sup\limits_{k\leqslant n} X_k \leqslant \lambda\right\}} X_n dP,\end{aligned}$$

此即 (1.11) 式.

又因为 $\left\{\sup\limits_{k\leqslant n} X_k > \lambda\right\} \subset \left\{\sup\limits_{k\leqslant n} X_k > \lambda\right\} \bigcup \left\{\inf\limits_{k\leqslant n} X_k < -\lambda\right\}$,

$$\boldsymbol{E}(X_n^-) = \int_{\{X_n<0\}} -X_n dP \geqslant \int_A -X_n dP$$

对一切 $A \in \mathscr{F}$ 成立, 所以把 (1.11) 与 (1.12) 式相加即得 (1.13) 式.

若 $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是非负下鞅, 则 $\{-X_n, \mathscr{F}_n, n \geqslant 0\}$ 是上鞅, 故

$$\begin{aligned}\lambda P(X_n^* > \lambda) \;&=\; \lambda P\left(\sup_{k\leqslant n} X_k > \lambda\right) = \lambda P\left(\inf_{k\leqslant n}(-X_k) < -\lambda\right) \\ &\overset{(1.12)}{\leqslant} \int_{\left\{\inf\limits_{k\leqslant n}(-X_k)<-\lambda\right\}} -(-X_n)dP = \int_{\left\{\sup\limits_{k\leqslant n} X_k>\lambda\right\}} |X_n|dP.\end{aligned}$$

(1.13)′ 式得证. 由 (1.13)′ 式对 $n \to \infty$ 取极限可得 (1.13)″ 式. 若 $\{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅, 用命题 1.4 亦可证 (1.13)′、(1.13)″ 式成立.

**定理 1.3 (Kolmogorov 不等式)** 设 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是鞅,$X_n^*,X^*$ 定义如前, 则

$$(1)\ \lambda^2P(X_n^*>\lambda)\leqslant \boldsymbol{E}(X_n^2)(n\geqslant 0,\lambda>0); \tag{1.14}$$

$$(2)\ \lambda^2P(X^*>\lambda)\leqslant \sup_{n\geqslant 0}\boldsymbol{E}(X_n^2)(\lambda>0). \tag{1.15}$$

**证** (1) 不妨设 $\boldsymbol{E}(X_n^2)<\infty$, 否则, (1.14) 式显然成立. 由于 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是鞅, 所以

$$X_k^2=[\boldsymbol{E}(X_n|\mathscr{F}_k)]^2\quad (k\leqslant n).$$

再用 Jensen 不等式得

$$\begin{aligned}\boldsymbol{E}(X_k^2)&=\boldsymbol{E}([\boldsymbol{E}(X_n|\mathscr{F}_k)]^2)\leqslant \boldsymbol{E}(\boldsymbol{E}(X_n^2|\mathscr{F}_k))\\&=\boldsymbol{E}(X_n^2)<\infty\quad (k\leqslant n).\end{aligned}$$

所以由命题 1.3 知 $\{-X_k^2,\mathscr{F}_k,k=0,1,\cdots,n\}$ 为上鞅, 对此上鞅及 $\lambda^2$ 用 (1.11) 式 (即在 (1.12) 式中以 $-X_k^2$ 代 $X_k$) 得

$$\begin{aligned}\lambda^2P(X_n^*>\lambda)&=\lambda^2P\left(\bigcup_{k\leqslant n}\{|X_k|>\lambda\}\right)\\&=\lambda^2P\left(\bigcup_{k\leqslant n}\{-X_k^2<-\lambda^2\}\right)\\&=\lambda^2P\left(\inf_{k\leqslant n}(-X_k^2)<-\lambda^2\right)\\&\leqslant\int_{\left\{\inf\limits_{k\leqslant n}(-X_k^2)<-\lambda^2\right\}}-(-X_n^2)dP\\&\leqslant \boldsymbol{E}(X_n^2).\end{aligned}$$

即 (1.14) 式成立.

(2) 在 (1.14) 式中令 $n\to\infty$ 取上极限即得 (1.15) 式.

**定理 1.4 (Doob 不等式)** 设 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为鞅或非负下鞅, $X_n^*,X^*$ 如前, 则

$$(1)\ \boldsymbol{E}(X^*)\leqslant\frac{e}{e-1}\left(1+\sup_{n\geqslant 0}\boldsymbol{E}(|X_n|\log^+|X_n|\right); \tag{1.16}$$

$$(2)\ \|X^*\|_p\leqslant q\|X\|_p\left(p>1,q>1,\frac{1}{p}+\frac{1}{q}=1\right), \tag{1.17}$$

其中 $\|X^*\|_p = \left(\int_\Omega |X^*|^p dP\right)^{\frac{1}{p}}$ (未必小于 $\infty$),

$$\|X\|_p = \sup_{n\geqslant 0} \|X_n\|_p = \sup_{n\geqslant 0} \left(\int_\Omega |X_n|^p dP\right)^{\frac{1}{p}} \quad (\text{未必小于 } \infty),$$

$$f^+ = f \vee 0.$$

**证**　(1) 任取 $n \geqslant 0, \lambda > 0$, 令 $F_n(\lambda) = P(X_n^* > \lambda)$. 由 (1.11) 式得

$$F_n(\lambda) \leqslant \frac{1}{\lambda} \int_{|X_n^* > \lambda|} |X_n| dP \tag{1.18}$$

设 $\Psi(\lambda)$ 为 $\mathbf{R}$ 上的非降右连续函数且 $\Psi(0) = 0$, 则由 $F_n$ 的定义并用分部积分公式及 Fubini 定理可得

$$\begin{aligned}
\boldsymbol{E}(\Psi(X_n^*)) &= -\int_{(0,\infty)} \Psi(\lambda) dF_n(\lambda) \\
&= \int_{(0,\infty)} F_n(\lambda) d\Psi(\lambda) - (\Psi(\infty)F_n(\infty) - \Psi(0)F_n(0)) \\
&\leqslant \int_{(0,\infty)} F_n(\lambda) d\Psi(\lambda) \\
&\overset{(1.18)}{\leqslant} \int_{(0,\infty)} \left(\frac{1}{\lambda} \int_{\{X_n^* > \lambda\}} |X_n| dP\right) d\Psi(\lambda) \\
&= \boldsymbol{E}\left(|X_n| \int_{(0,X_n^*)} \frac{1}{\lambda} d\Psi(\lambda)\right).
\end{aligned} \tag{1.19}$$

在 (1.19) 式中取 $\Psi(\lambda) = (\lambda - 1)^+$, 若注意 $\Psi(\lambda)$ 产生的 Lebesgue-Stieltjes 测度在 $(-\infty, 1]$ 之测度为 0, 则得

$$\begin{aligned}
\boldsymbol{E}((X_n^* - 1)) &\leqslant \boldsymbol{E}((X_n^* - 1)^+) \\
&\overset{(1.19)}{\leqslant} \boldsymbol{E}\left(|X_n| \int_{(0,X_n^*)} \frac{1}{\lambda} d\Psi(\lambda); X_n^* \leqslant 1\right) \\
&\quad + \boldsymbol{E}\left(|X_n| \int_{(0,X_n^*)} \frac{1}{\lambda} d\Psi(\lambda); X_n^* > 1\right) \\
&= 0 + \boldsymbol{E}\left(|X_n| \int_{(1,X_n^*)} \frac{1}{\lambda} d\lambda; X_n^* > 1\right) \\
&= \boldsymbol{E}(|X_n| \log X_n^*; X_n^* > 1) \\
&= \boldsymbol{E}\left(|X_n| \log^+ X_n^*\right).
\end{aligned} \tag{1.20}$$

由于 $\log x \leqslant \dfrac{x}{e} (x \geqslant 0)$, 故对任意 $a \geqslant 0, b \geqslant 0$, 有

$$a \log^+ b \leqslant a \log^+ a + \frac{b}{e}, \tag{1.21}$$

从而

$$E(|X_n|\log^+ X_n^*) \leqslant E(|X_n|\log^+|X_n|) + \frac{1}{e}E(X_n^*).$$

代入 (1.20) 式得

$$E(X_n^*) \leqslant 1 + E(|X_n|\log^+|X_n|) + \frac{1}{e}E(X_n^*),$$

即

$$E(X_n^*) \leqslant \frac{e}{e-1}(1 + E(|X_n|\log^+|X_n|)).$$

但 $X_n^* \uparrow X^*(n\uparrow\infty)$, 所以在上式中令 $n\to\infty$ 并应用 Fatou 引理可得 (1.16) 式.

(2) 在 (1.19) 式中取 $\Psi(\lambda)=\lambda^p$, 则得

$$\begin{aligned}E((X_n^*)^p) &\leqslant E\left(|X_n|\int_{(0,X_n^*)} p\cdot\lambda^{p-2}d\lambda\right)\\ &= \frac{p}{p-1}E(|X_n|(X_n^*)^{p-1})\\ &= qE(|X_n|(X_n^*)^{p-1}).\end{aligned}$$

再用 Hölder 不等式得

$$\begin{aligned}E((X_n^*)^p) &\leqslant q(E(|X_n|^p)^{\frac{1}{p}}\cdot E((X_n^*)^{q(p-1)})^{\frac{1}{q}})\\ &= q(E(|X_n|^p)^{\frac{1}{p}}E((X_n^*)^p)^{\frac{1}{q}}).\end{aligned}\tag{1.22}$$

为证 (1.17) 式, 不妨设 $\|X\|_p<\infty$. 于是

$$\|X_n^*\|_p \leqslant \left\|\sum_{k=0}^n |X_k|\right\|_p \leqslant \sum_{k=0}^n \|X_k\|_p < \infty.\tag{1.23}$$

若 $\|X_n^*\|_p\equiv 0$ (一切 $n\geqslant 0$), 则由 $X_n^*\uparrow X^*$ 得 $\|X^*\|_p=0$, 从而 (1.17) 式成立. 若存在一个 $n_0\geqslant 0$, 使 $\|X_{n_0}^*\|_p>0$, 则 $\|X_n^*\|_p>0$ (对一切 $n\geqslant n_0$). 所以由 (1.23) 式知 $\|X_n^*\|_p$ 是正实数 $(n\geqslant n_0)$, 故在 (1.22) 式两边除以 $E((X_n^*)^p)^{\frac{1}{q}}=\|X_n^*\|_p^{\frac{p}{q}}$ 可得

$$\|X_n^*\|_p \leqslant q\|X_n\|_p.$$

而 $X_n^*\uparrow X^*$, 所以在上式中令 $n\to\infty$ 取极限即得 (1.17) 式.

## §2 鞅的收敛定理

在这一节中, 我们研究鞅的收敛定理. 首先给出几个引理.

**引理 2.1 (上穿不等式)** 设 $x=\{x_n, n=0,1,\cdots\}$ 为广义实数列 (即 $x_n$ 可为实数亦可为 $\infty$), $a$ 和 $b$ 为二个实数, $a<b$. 令

$$u_0(x)=1, \quad 对\ n\geqslant 0\ 再令$$

$$u_{n+1}(x)=\begin{cases}1, & 若\ x_n>b,\\ u_n(x), & 若\ x_n\in[a,b],\\ 0, & 若\ x_n<a.\end{cases} \tag{2.1}$$

再设 $g_n^{(a,b)}(x)$ 是上穿函数, 即

$$g_n^{(a,b)}(x)=\begin{cases}1, & 当\ (x_0,x_1,\cdots)\ 在时刻\ n\ 上穿\ [a,b],\\ 0, & 反之, \qquad (n\geqslant 0)\end{cases} \tag{2.2}$$

所谓 $(x_0,x_1,\cdots)$ 在时刻 $n$ 上穿 $[a,b]$, 即是 $x_n>b$, 而且存在正整数 $m>0$, 使 $x_{n-m}<a$, 并且 $x_k\in[a,b]$ (当 $n>k>n-m$). 显然 $g_0^{(a,b)}(x)=0$. 再令

$$U_n^{(a,b)}(x)=\sum_{k=0}^{n} g_k^{(a,b)}(x)=\sum_{k=1}^{n} g_k^{(a,b)}(x) \tag{2.3}$$

为 $x=(x_0,x_1,\cdots)$ 到时刻 $n$ 为止上穿 $[a,b]$ 的次数, 亦即 $x_0,x_1,\cdots,x_n$ 上穿 $[a,b]$ 的次数, 则有如下结果:

$$(b-a)U_n^{(a,b)}(x)\leqslant \sum_{k=1}^{n}(x_k-a)(u_{k+1}(x)-u_k(x)) \quad (n\geqslant 1). \tag{2.4}$$

证明甚易, 读者可作为习题验证之.

**引理 2.2 (Doob 下鞅上穿不等式)** 设 $\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为下鞅, 记 $X=(X_0,X_1,\cdots)$. 定义 $u_k(X), g_n^{(a,b)}(X)$ 和 $U_n^{(a,b)}(X)$ 如引理 2.1. 再令

$$U^{(a,b)}(X)=\lim_{n\to\infty} U_n^{(a,b)}(X)=\sum_{k=0}^{\infty} g_k^{(a,b)}(X)$$

是 $X$ 上穿 $[a,b]$ 的总次数, 则

$$\boldsymbol{E}(U_n^{(a,b)}(X))\leqslant \frac{\boldsymbol{E}(|X_n|)+|a|}{b-a}; \tag{2.5}$$

$$\boldsymbol{E}(U^{(a,b)}(X))\leqslant \frac{\sup\limits_{n\geqslant 0}\boldsymbol{E}(|X_n|)+|a|}{(b-a)}. \tag{2.6}$$

**证** 利用引理 2.1, 证明此引理甚易. 读者可作为习题验证之.

**引理 2.3** 设 $\xi$ 和 $\xi_n(n\geqslant 0)$ 皆为概率空间 $(\Omega,\mathscr{F},P)$ 上的随机变量, 它们的数学期望都存在, 则

$$“\xi_n \xrightarrow{L'} \xi \Leftrightarrow \{\xi_n, n\geqslant 0\} \text{ 一致可积且 } \xi_n \xrightarrow{P} \xi”.$$

所谓 $\{\xi_n, n\geqslant 0\}$ 一致可积, 是指任给 $\varepsilon>0$, 总存在正数 $K$, 使

$$\sup_{n\geqslant 0}\int_{\{|\xi_n|\geqslant K\}}|\xi_n|dP\leqslant\varepsilon.$$

**证** 此系测度论中常用之结果, 读者不妨作为习题复习一下, 补证之.

**定理 2.1 (Doob 鞅的收敛定理)** 设 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的下鞅或上鞅, $\|X\|_1\overset{\text{def.}}{=}\sup\limits_{n\geqslant 0}\|X_n\|_1\overset{\text{def.}}{=}\sup\limits_{n\geqslant 0}\boldsymbol{E}(|X_n|)<\infty$ (即 $X$ 是 $L^1$ 有界 的下鞅或上鞅), 则

$$\lim_{n\to\infty}X_n=X_\infty,\quad \text{a.s.},$$
$$\boldsymbol{E}(|X_\infty|)\leqslant\|X\|_1<\infty.$$

**证** 设 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 是下鞅. 令 $U^{(a,b)}(X)$ 的定义如引理 1.2. 则有

$$\boldsymbol{E}(U^{(a,b)}(X))\leqslant\frac{\|X\|_1+|a|}{b-a}<\infty,$$

从而

$$\begin{aligned}0&=P(U^{(a,b)}(X)=\infty)\\&\geqslant P\left(\liminf_{n\to\infty}X_n<a<b<\limsup_{n\to\infty}X_n\right).\end{aligned}\tag{2.7}$$

令 $Q$ 为有理数集, 则

$$\begin{aligned}\left\{\lim_{n\to\infty}X_n\text{ 不存在}\right\}^{①}&\subset\left\{\liminf_{n\to\infty}X_n<\limsup_{n\to\infty}X_n\right\}\\&\subset\bigcup_{\substack{a<b\\a,b\in Q}}\left\{\liminf_{n\to\infty}X_n<a<b<\limsup_{n\to\infty}X_n\right\}.\end{aligned}\tag{2.8}$$

由 (2.7)、(2.8) 式得

$$\lim_{n\to\infty}X_n=X_\infty,\quad\text{a.s.}.$$

再用 Fatou 引理可得

$$\begin{aligned}\boldsymbol{E}(|X_\infty|)&=\boldsymbol{E}\left(\lim_{n\to\infty}|X_n|\right)\leqslant\liminf_{n\to\infty}\boldsymbol{E}(|X_n|)\\&\leqslant\sup_{n\geqslant 0}\boldsymbol{E}(|X_n|)=\|X\|_1<\infty.\end{aligned}$$

① 此处及今后所言的 $\lim\limits_{n\to\infty}X_n$ 存在 (或收敛), 均指它可以为 $\infty$. 若 $\lim\limits_{n\to\infty}X_n$ 存在且为有穷, 则 “有穷” 二字要强调指出.

**定理 2.2** 设 $X=\{X_n,\mathscr{F}_n,n\geqslant 0\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的鞅 (相应地, 为上鞅), $\{X_n\}$ 一致可积, 则存在随机变量 $X_\infty, \boldsymbol{E}(|X_\infty|)<\infty$, 使

$$\lim_{n\to\infty} X_n = X_\infty, \quad \text{a.s. 及 } L^1, \tag{2.9}$$

且对一切 $n\geqslant 0$, 有

$$\boldsymbol{E}(X_\infty|\mathscr{F}_n)=X_n \quad (\text{相应地}, \leqslant X_n). \tag{2.10}$$

**证** 由于 $\{X_n\}$ 是一致可积鞅, 所以

$$\sup_{n\geqslant 0}\boldsymbol{E}(|X_n|)<\infty,$$

由定理 2.1, 存在随机变量 $X_\infty$, 使

$$\lim_{n\to\infty} X_n = X_\infty, \quad \text{a.s.},$$
$$\boldsymbol{E}(|X_\infty|)<\infty.$$

再用引理 1.3 得

$$\lim_{n\to\infty} X_n = X_\infty, \quad L^1..$$

最后, 推证 (2.10) 式成立. 由于 $X_n\in\mathscr{F}_n$, 只需证明

$$\boldsymbol{E}(X_n;\Lambda_n)=\boldsymbol{E}(X_\infty;\Lambda_n) \quad (\Lambda_n\in\mathscr{F}_n). \tag{2.11}$$

事实上, 任取 $\Lambda_n\in\mathscr{F}_n$, 有

$$\lim_{m\to\infty}\mathbf{1}_{\Lambda_n}X_{m+n}=\mathbf{1}_{\Lambda_n}X_\infty, [L^1],$$
$$\lim_{m\to\infty}\boldsymbol{E}(\mathbf{1}_{\Lambda_n}X_{m+n})=\boldsymbol{E}(\mathbf{1}_{\Lambda_n}X_\infty).$$

但是

$$\boldsymbol{E}(\mathbf{1}_{\Lambda_n}X_{m+n})=\boldsymbol{E}(X_n\mathbf{1}_{\Lambda_n}),$$

在上式中令 $m\to\infty$ 得

$$\boldsymbol{E}(\mathbf{1}_{\Lambda_n}X_\infty)=\boldsymbol{E}(\mathbf{1}_{\Lambda_n}X_n).$$

即 (2.11) 式成立. 定理证毕.

关于连续时间参数的鞅, 亦有类似定理 1.2~1.4 及定理 2.1, 2.2 的结果.

设 $x:[0,\infty)\mapsto[-\infty,\infty], A=\{t_1,\cdots,t_n\}\subset[0,\infty)$, 将 $A$ 中元素依大小顺序排列: $t_{k_1}<t_{k_2}<\cdots<t_{k_n}$. 令 $U_A^{(a,b)}(x)$ 为 $x(t_1),x(t_2),\cdots,x(t_k)$ 上穿 $[a,b]$ 的次数 (定义见 (2.3) 式). 对 $[0,\infty)$ 中任一子集 $B$, 定义

$$U_B^{(a,b)}(x)=\sup_{\substack{A\subset B\\ A\ \text{为有限集}}} U_A^{(a,b)}(x).$$

若 $B=\{t_1,t_2,\cdots\}$ 为可数集, $A_n=\{t_1,t_2,\cdots,t_n\}$, 则显然有

$$U_B^{(a,b)}(x)=\lim_{n\to\infty}U_{A_n}^{(a,b)}(x).$$

**定理 1.2′** 设 $\{X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的下鞅, $D$ 为 $[0,\infty)$ 上任一可数稠子集, 则对任何 $r<s(r,s\in[0,\infty)),a<b(a,b\in\mathbf{R})$ 及 $\lambda>0$, 有

$$(1)\qquad \lambda P\left(\sup_{t\in D\cap[r,s]}|X_t|>\lambda\right)\leqslant -\boldsymbol{E}(X_r)+2\boldsymbol{E}(X_s^+);\tag{2.12}$$

$$(2)\qquad \boldsymbol{E}\left(U_{D\cap[r,s]}^{(a,b)}(X)\right)\leqslant\frac{\boldsymbol{E}(|X_s|)+|a|}{-a+b}.\tag{2.13}$$

若 $\{X_t\}$ 的几乎所有的轨道右连续, 则 (2.12)、(2.13) 式中的 $D\bigcap[r,s]$ 可以代之以 $[r,s]$.

**证** 设 $D\bigcap[r,s]=\{t_0,t_1,\cdots\}$, 令 $A_n=\{t_0,\cdots,t_n\}$, 而 $\{-X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 为上鞅, 则由 (1.13) 式得

$$\begin{aligned}\lambda P\left(\sup_{k\leqslant n}|X_{t_k}|>\lambda\right)&\leqslant \boldsymbol{E}(-X_{t_0})+2\boldsymbol{E}((-X_{t_n})^-)\\&=-\boldsymbol{E}(X_{t_0})+2\boldsymbol{E}(X_{t_n}^+),\\&\leqslant -\boldsymbol{E}(X_r)+2\boldsymbol{E}(X_s^+)\quad(\text{因 } r\leqslant t_0\leqslant t_n\leqslant s).\end{aligned}$$

令 $n\to\infty$ 即得 (2.12) 式.

仿之可证 (2.13) 式, 只不过在使用 (1.13) 式的地方将其改用 (2.5) 式罢了.

定理的最后一个论断是明显的. 定理证毕.

**定理 1.3′ (Kolmogorov 不等式)** 设 $\{X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的鞅, 对 a.s. 的 $\omega,X(\cdot,\omega)$ 右连续, $X^*=\sup\limits_{t\in[0,\infty)}|X_t|$, 则

$$\lambda^2P(X^*>\lambda)\leqslant\sup_{t\in[0,\infty]}\boldsymbol{E}(X_t^2)\quad(\lambda>0).\tag{2.14}$$

**定理 1.4′ (Doob 不等式)** 设 $\{X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上的鞅或非负下鞅, 对 a.s. 的 $\omega,X(\cdot,\omega)$ 右连续, $X^*=\sup\limits_{t\in[0,\infty]}|X_t|$, 则

$$(1)\qquad \boldsymbol{E}(X^*)\leqslant\frac{e}{e-1}\left(1+\sup_{t\in[0,\infty)}\boldsymbol{E}(|X_t|\log^+|X_t|)\right),\tag{2.15}$$

(2) $\|X^*\|_p\leqslant q\|X\|_p=q\sup\limits_{t\in[0,\infty)}\|X_t\|_p$

$$\left(p>1,q>1,\frac{1}{p}+\frac{1}{q}=1\right).\tag{2.16}$$

**定理 2.1′ (Doob 收敛定理)** 设 $X=\{X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的下鞅或上鞅, $\|X\|_1\equiv\sup\limits_{t\in[0,\infty)}\boldsymbol{E}(|X_t|)<\infty$, 对 a.s. 的 $\omega, X(\cdot,\omega)$ 右连续, 则存在随机变量 $X_\infty$, 使

$$\lim_{t\to\infty}X_t=X_\infty,\quad \text{a.s.},\tag{2.17}$$

$$\boldsymbol{E}(|X_\infty|)\leqslant\|X\|_1<\infty.\tag{2.18}$$

**定理 2.2′** 设 $X=\{X_t,\mathscr{F}_t,t\in[0,\infty)\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的鞅 (相应地, 上鞅), $\{X_t\}$ 一致可积, 对 a.s. 的 $\omega, X(\cdot,\omega)$ 右连续, 则存在随机变量 $X_\infty, \boldsymbol{E}(|X_\infty|)<\infty$, 使

$$\lim_{t\to\infty}X_t=X_\infty,\quad \text{a.s. 及 } L^1_\bullet,\tag{2.19}$$

且对一切 $t\in[0,\infty)$, 有

$$\boldsymbol{E}(X_\infty|\mathscr{F}_t)=X_t\quad (\text{相应地}, \leqslant X_t).\tag{2.20}$$

## *§3 鞅的 Doob 停时理论

"停时" 这一概念, 在随机过程的理论及应用中, 是一个十分重要的概念, 特别是在 "强 Markov" 理论及鞅论中更为重要. 限于篇幅, 本书未涉及强 Markov 过程.

**定义 3.1** 设 $(\Omega,\mathscr{F})$ 是一个可测空间, $T=[0,\infty], \{\mathscr{F}_t,t\in[0,\infty]\}$ 是 $\mathscr{F}$ 中的一族单调上升的子 $\sigma$ 代数 (即 $s\leqslant t\Rightarrow\mathscr{F}_s\subset\mathscr{F}_t$) 有时简记之为 $\{\mathscr{F}_t\}$. 再记

$$\mathscr{F}_{t+}\overset{\text{def.}}{=\!=}\bigcap_{s>t}\mathscr{F}_s\quad(0\leqslant t<\infty),$$

$$\mathscr{F}_{t-}\overset{\text{def.}}{=\!=}\bigvee_{s<t}\mathscr{F}_s\overset{\text{def.}}{=\!=}\sigma\left(\bigcup_{s<t}\mathscr{F}_s\right)\quad(0<t<\infty).$$

显然 $\{\mathscr{F}_{t+},t\in[0,\infty]\}$ 和 $\{\mathscr{F}_{t-},0<t\leqslant\infty\}$ 都是 $\mathscr{F}$ 中的单调上升的子 $\sigma$ 代数族, $\mathscr{F}_{\infty-}\overset{\text{def.}}{=\!=}\mathscr{F}_{\infty+}\overset{\text{def.}}{=\!=}\mathscr{F}_\infty\overset{\text{def.}}{=\!=}\mathscr{F}, \mathscr{F}_{0-}\overset{\text{def.}}{=\!=}\mathscr{F}_0$. 此外, 还有 $\mathscr{F}_{t-}\subset\mathscr{F}_t\subset\mathscr{F}_{t+}(\forall t\in[0,\infty])$

称 $\mathscr{F}$ 中的子 $\sigma$ 代数族 $\{\mathscr{F}_t,t\in[0,\infty]\}$ **是右连续的**, 如果 $\mathscr{F}_{t+}=\mathscr{F}_t, (\forall t\in[0,\infty])$.

显然,$\{\mathscr{F}_{t+},t\in[0,\infty]\}$ 是右连续的.

**定义 3.2** 设 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 如定义 3.1. 称映射 $\tau:\Omega\mapsto[0,\infty]$ 是 $\{\mathscr{F}_t, t\in[0,\infty]\}$ **停时** (简称停时), 如果

$$\{\tau\leqslant t\}\in\mathscr{F}_t\quad(\forall t\in[0,\infty)).$$

若注意: $\{\tau=\infty\}=\Omega-\{\tau<\infty\}\in\mathscr{F}=\mathscr{F}_\infty$ 可知:

$\tau$ 是 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 停时等价于

$$\{\tau\leqslant t\}\in\mathscr{F}_t\quad(\forall t\in[0,\infty]).$$

对停时 $\tau$, 显然还等价于

$$\{\tau> t\}\in\mathscr{F}_t\quad(\forall t\in[0,\infty)).$$

**命题 3.1** 对 $\mathscr{F}$ 中任何一族单调上升子 $\sigma$ 代数 $\{\mathscr{F}_t, t\in[0,\infty]\}$, 总有:

(1) $\tau$ 是 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 停时 $\Rightarrow\tau$ 是 $\{\mathscr{F}_{t+}, t\in[0,\infty]\}$ 停时;

(2) $\tau$ 是 $\{\mathscr{F}_{t+}, t\in[0,\infty]\}$ 停时的充分必要条件是

$$\{\tau< t\}\in\mathscr{F}_t\quad(\forall t\in[0,\infty]).$$

证明甚易, 读者可作为习题验证之.

**命题 3.2** 若 $\tau_1$ 和 $\tau_2$ 都是 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 停时, 则 $\tau_1+\tau_2, \max\{\tau_1,\tau_2\}$, $\min\{\tau_1,\tau_2\}$ 亦然.

证明甚易, 读者可作为习题验证之.

**命题 3.3** 若 $\{\tau_n, n\geqslant 0\}$ 是一列 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 停时, 则 $\sup\limits_{n\geqslant 0}\tau_n$ 亦然; 若还有 $\mathscr{F}_t=\mathscr{F}_{t+}(\forall t\in[0,\infty])$, 则 $\inf\limits_{n\geqslant 0}\tau_n, \liminf\limits_{n\to\infty}\tau_n$ 和 $\limsup\limits_{n\to\infty}\tau_n$ 亦然.

证明甚易. 读者可作为习题验证之.

**定义 3.3** 若 $\tau$ 是 $\{\mathscr{F}_t\}$ 停时, 称

$$\mathscr{F}_\tau \overset{\text{def.}}{=\!=} \{A\in\mathscr{F}_\infty : A\bigcap\{\tau\leqslant t\}\in\mathscr{F}_t,\quad \forall t\in[0,\infty]\}$$

为 $\tau$ **前 $\sigma$ 代数**.

显然 $\mathscr{F}_\tau$ 是 $\sigma$ 代数, 若 $\tau\equiv a$ 是常数, 则 $\mathscr{F}_\tau=\mathscr{F}_\alpha\cdot\mathscr{F}_\tau$ 可理解为某物理过程到时刻 $\tau$ 为止的全部信息. “停时” $\tau$ 的含义意味着当我们仅仅知道到 $t$ 为止的信息后, 就能断言 $\tau$ 是否大于 $t$.

若 $\tau$ 是 $\{\mathscr{F}_{t+}\}$ 停时, 记

$$\mathscr{F}_{\tau+} \overset{\text{def.}}{=\!=} \{A\in\mathscr{F}_\infty : A\bigcap\{\tau\leqslant t\}\in\mathscr{F}_{t+},\quad \forall t\in[0,\infty]\}$$

易证 $\mathscr{F}_{\tau+}$ 是 $\sigma$ 代数, 而且

$$\mathscr{F}_{\tau+} = \{A \in \mathscr{F}_\infty : A \bigcap \{\tau < t\} \in \mathscr{F}_t, \quad \forall t \in [0,\infty]\}.$$

**命题 3.4** 任给单增 $\sigma$ 代数族 $\{\mathscr{F}_t\}$, 有

(1) $\tau$ 是 $\{\mathscr{F}_t\}$ 停时 $\Rightarrow \tau \in \mathscr{F}_\tau$ (即 $\tau$ 关于 $\mathscr{F}_\tau$ 可测);

(2) $\tau_1, \tau_2$ 皆为 $\{\mathscr{F}_t\}$ 停时, $\tau_1 \leqslant \tau_2 \Rightarrow \mathscr{F}_{\tau_1} \subset \mathscr{F}_{\tau_2}$;

(3) $\tau_n$ 是 $\{\mathscr{F}_t\}$ 停时 $(n \geqslant 1)$, $\{\mathscr{F}_t\}$ 右连续, $\tau = \inf \tau_n \Rightarrow \mathscr{F}_\tau = \bigcap\limits_{n=1}^{\infty} \mathscr{F}_{\tau_n}$.

证明甚易, 读者可作为习题验证之.

**命题 3.5** 设 $\tau_1$ 和 $\tau_2$ 皆为 $\{\mathscr{F}_t, t \in [0,\infty]\}$ 停时, 则 $\{\tau_1 < \tau_2\}$, $\{\tau_1 \leqslant \tau_2\}$ 和 $\{\tau_1 = \tau_2\}$ 皆属于 $\mathscr{F}_{\tau_1} \bigcap \mathscr{F}_{\tau_2}$.

证明甚易, 读者可作为习题验证之.

注意: 前面我们讨论的是取值于 $[0,\infty]$ 的 $\{\mathscr{F}_t, t \in [0,\infty]\}$ 停时 $\tau$.

若 $(\Omega, \mathscr{F})$ 是一可测空间, $\{\mathscr{F}_n, n = 0, 1, \cdots, \infty\}$ 是一族 $\mathscr{F}$ 中的单调上升的子 $\sigma$ 代数, 我们可以类似地定义取值于 $\{0, 1, \cdots, \infty\}$ 关于 $\{\mathscr{F}_n, n = 0, 1, \cdots, \infty\}$ 的离散的停时, 而且有类似的定义及命题 (当然涉及 $\mathscr{F}_{t+}$ 的定义及命题是不足道的). 例如

**定义 3.2′** 设 $\{\mathscr{F}_n, n = 0, 1, \cdots, \infty\}$ 是 $\mathscr{F}$ 中一族单调上升的子 $\sigma$ 代数, $\tau : \Omega \mapsto \{0, 1, \cdots, \infty\}$. 称 $\tau$ 是 $\{\mathscr{F}_n\}$ 停时, 如果对任何非负整数 $n$, 有

$$\{\tau \leqslant n\} \in \mathscr{F}_n.$$

**定义 3.3′** 若 $\tau$ 是 $\{\mathscr{F}_n\}$ 停时, 亦称

$$\mathscr{F}_\tau \stackrel{\text{def.}}{=\!=} \{A \in \mathscr{F}_\infty : A \bigcap \{\tau \leqslant n\} \in \mathscr{F}_n$$

对任何 $n = 0, 1, \cdots, \infty$ 成立 $\}$ 为 $\tau$ 前 $\sigma$ 代数.

除涉及 $\mathscr{F}_{t+}$ 的内容外, 命题 3.2 到命题 3.5 的结论依然成立. 读者可作为习题表述并证明之.

现在我们研究鞅的 Doob 停时理论, 主要研究离散时间鞅的 Doob 停时理论.

先从离散时间鞅讨论起, 设 $X = \{X_n, \mathscr{F}_n, n = 0, 1, \cdots, \infty\}$ 为概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $\tau$ 为 $\{\mathscr{F}_n\}$ 停时. 令

$$X_n^\tau = X_{\tau \wedge n}; \quad (\tau \wedge n \stackrel{\text{def.}}{=\!=} \min(\tau, n)), \tag{3.1}$$

$$X^\tau = \{X_n^\tau, \mathscr{F}_n, n = 0, 1, \cdots\}, \tag{3.2}$$

$X^\tau$ ($X^\tau$ 称为 $X$ 的 $\tau$ 前过程) 是否为鞅? 若 $\tau_k$ 是 $\{\mathscr{F}_n\}$ 停时, $\tau_0 \leqslant \tau_1 \leqslant \cdots$, $\mathscr{F}_{\tau_k}$ 是 $\tau_k$ 前 $\sigma$ 代数 (见定义 3.3), $\{X_{\tau_n}, \mathscr{F}_{\tau_n}, n = 0, 1, \cdots\}$ 是否为鞅? 回答都是肯定的.

**定理 3.1** 若 $X = \{X_n, \mathscr{F}_n, n = 0, 1, \cdots\}$ 为概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅 (相应地, 上鞅), $\tau$ 为 $\{\mathscr{F}_n\}$ 停时, 则 $X^\tau$ 是鞅 (相应地, 上鞅), 且 $\boldsymbol{E}(X_0) = \boldsymbol{E}(X_{\tau\wedge n}) = \boldsymbol{E}(X_n)$ (相应地, $\boldsymbol{E}(X_0) \geqslant \boldsymbol{E}(X_{\tau\wedge n}) \geqslant \boldsymbol{E}(X_n)$), $n \geqslant 0$.

若 $X$ 是 $L^1$ 有界的鞅或非负下鞅, 则

$$\|X_\tau\|_1 \leqslant \|X^\tau\|_1 \leqslant \|X\|_1. \tag{3.3}$$

($X_\infty$ 定义为 $\lim\limits_{n\to\infty} X_n$, $\|X_\tau\|_1$ 是 $X_\tau$ 的 $L^1$ 范数, $\|X^\tau\|_1 \overset{\text{def.}}{=\!=} \sup\limits_{n\geqslant 0} \|X_n^\tau\|_1$.)

若 $X$ 是一致可积的鞅 (相应地, 上鞅), 则 $X_\tau$ 亦然.

**证** 令 $D_n = X_n - X_{n-1} (n \geqslant 0, X_{-1} \equiv 0)$, 因 $X$ 是鞅, $\tau$ 是 $\{\mathscr{F}_n\}$ 停时, 故

$$\mathbf{1}_{\{\tau\geqslant n\}} \in \mathscr{F}_{n-1}, \boldsymbol{E}(D_n|\mathscr{F}_{n-1}) = 0 (n \geqslant 1), \tag{3.4}$$

但是

$$\begin{aligned} X_{n\wedge\tau} &= \sum_{k=0}^{n} X_k \mathbf{1}_{\{\tau=k\}} + X_n \mathbf{1}_{\{\tau>n\}} \\ &= \sum_{k=0}^{n}\sum_{j=0}^{k} D_j \mathbf{1}_{\{\tau=k\}} + \sum_{j=0}^{n} D_j \mathbf{1}_{\{\tau>n\}} \\ &= \sum_{j=0}^{n} D_j \mathbf{1}_{\{n\geqslant\tau\geqslant j\}} + \sum_{j=0}^{n} D_j \mathbf{1}_{\{\tau>n\}} \\ &= \sum_{j=0}^{n} \mathbf{1}_{\{\tau\geqslant j\}} D_j. \end{aligned} \tag{3.5}$$

由 (3.5) 式得 $\boldsymbol{E}(|X_{n\wedge\tau}|) < \infty$, 且

$$\begin{aligned} \boldsymbol{E}(X_n) - \boldsymbol{E}(X_{n\wedge\tau}) &\overset{(3.5)}{=\!=} \sum_{j=0}^{n} \boldsymbol{E}(\mathbf{1}_{\{\tau<j\}} D_j) \\ &= \sum_{j=1}^{n} \boldsymbol{E}(\mathbf{1}_{\{\tau<j\}} D_j) \\ &= \sum_{j=1}^{n} \boldsymbol{E}(\mathbf{1}_{\{\tau<j\}} \boldsymbol{E}(D_j|\mathscr{F}_{j-1})) \overset{(3.4)}{=\!=} 0. \end{aligned}$$

显然 $\boldsymbol{E}(X_0) = \boldsymbol{E}(X_n)$.

下面证明 $X^\tau$ 是鞅, 任取 $B \in \mathscr{F}_n, \Lambda \in \mathscr{B}^1$, 有

$$\{X_{n\wedge\tau} \in \Lambda\} = \{X_n \in \Lambda : \tau \geqslant n\} \bigcup \left( \bigcup_{j=0}^{n-1} \{X_j \in \Lambda : \tau = j\} \right) \in \mathscr{F}_n \quad (n \geqslant 0).$$

$$\int_B X_{n\wedge\tau} dP = \sum_{j=0}^{n-1} \int_{B\cap\{\tau=j\}} X_j dP + \int_{B\cap\{\tau\geqslant n\}} X_n dP.$$

但是由 $X$ 是鞅, $B \bigcap \{\tau > n\} \in \mathscr{F}_n$, 得

$$\int_{B\cap\{\tau>n\}} X_n dP = \int_{B\cap\{\tau>n\}} X_{n+1} dP,$$

代入上式得

$$\begin{aligned}\int_B X_{n\wedge\tau} dP &= \sum_{j=0}^{n} \int_{B\cap\{\tau=j\}} X_j dP + \int_{B\cap\{\tau>n\}} X_{n+1} dP \\ &= \int_B X_{(n+1)\wedge\tau} dP \quad (n \geqslant 0).\end{aligned}$$

所以 $X^\tau$ 是鞅.

若 $X$ 是 $L^1$ 有界鞅或非负下鞅, 由于

$$X_\tau = \lim_{n\to\infty} X_{\tau\wedge n} \quad (\text{因 } X_\infty \equiv \lim_{n\to\infty} X_n),$$

所以由 Fatou 引理得

$$\begin{aligned}\|X_\tau\|_1 \equiv \boldsymbol{E}(|X_\tau|) &\leqslant \liminf_{n\to\infty} \boldsymbol{E}(|X_{\tau\wedge n}|) \\ &\leqslant \sup_{n\geqslant 0} \boldsymbol{E}(|X_{\tau\wedge n}|) \equiv \|X^\tau\|_1.\end{aligned}$$

但是 $|X| = \{|X_n|, \mathscr{F}_n, n = 0, 1, \cdots\}$ 是非负下鞅, $\{\tau = k\} \in \mathscr{F}_k$, 所以, 当 $k < n$ 时, 有

$$\int_{\{\tau=k\}} |X_k| dP \leqslant \int_{\{\tau=k\}} |X_n| dP.$$

因此

$$\begin{aligned}\boldsymbol{E}(|X_{\tau\wedge n}|) &= \sum_{k=0}^{n-1} \int_{\{\tau=k\}} |X_k| dP + \int_{\{\tau\geqslant n\}} |X_n| dP \\ &\leqslant \sum_{k=0}^{n-1} \int_{\{\tau=k\}} |X_n| dP + \int_{\{\tau\geqslant n\}} |X_n| dP \\ &= \boldsymbol{E}(|X_n|) \quad (n \geqslant 0).\end{aligned}$$

所以

$$\|X^\tau\|_1 \equiv \sup_{n\geqslant 0} \boldsymbol{E}(|X_{\tau\wedge n}|) \leqslant \sup_{n\geqslant 0} \boldsymbol{E}(|X_n|) \equiv \|X\|_1.$$

若 $X$ 是一致可积鞅, 如前所证, $X^\tau$ 必为鞅, 且 $\|X_\tau\|_1 < \infty$ (因为 $X$ 一致可积蕴涵了 $X$ 是 $L^1$ 有界的). 推证 $\{X_{n\wedge\tau} : n \geqslant n\}$ 是 $L^1$ 收敛, 果能如此, 用引理 1.3 得知 $\{X_{n\wedge\tau} : n \geqslant 0\}$ 必为一致可积的.

事实上, 当 $m < n$ 时, 有

$$\begin{aligned}\|X_{n\wedge\tau} - X_{m\wedge\tau}\|_1 \leqslant & \int_{\{\tau<m\}} |X_\tau - X_\tau| dP \\ & + \int_{\{m\leqslant t<n\}} |X_\tau - X_m| dP + \int_{\{\tau\geqslant n\}} |X_n - X_m| dP \\ \leqslant & \int_{\{m\leqslant\tau<\infty\}} |X_\tau - X_m| dP + \int_{\{\tau\geqslant n\}} |X_n - X_m| dP.\end{aligned}$$

但 $X$ 是一致可积鞅, 由定理 2.2 有

$$\lim_{n\to\infty} X_n = X_\infty, \quad \text{a.s. 和 } L^1.$$

$$\lim_{m,n\to\infty} \int_{\{\tau\geqslant n\}} |X_n - X_m| dP = 0.$$

令 $Y_m = \mathbf{1}_{\{m\leqslant\tau<\infty\}}|X_\tau - X_m|$, 则由

$$\lim_{n\to\infty} X_n = X_\infty, \quad \text{a.s.}$$

得

$$\lim_{n\to\infty} Y_m = 0, \quad \text{a.s.}$$

但是 $X_\tau$ 可积, $\{X_n : n \geqslant 0\}$ 一致可积, 所以 $\{Y_n : n \geqslant 0\}$ 是一致可积的. 因此, 由引理 1.3 得

$$\begin{aligned}\lim_{m\to\infty} \int_{\{m\leqslant\tau<\infty\}} |X_\tau - X_m| dP &= \lim_{m\to\infty} \int_\Omega Y_m dP \\ &= \int_\Omega \left(\lim_{m\to\infty} Y_m\right) dP = 0.\end{aligned}$$

故 $\lim\limits_{m,n\to\infty} \|X_{n\wedge\tau} - X_{m\wedge\tau}\|_1 = 0$, 即 $\{X_{n\wedge\tau} : n \geqslant 0\}$, 当 $n \to \infty$ 时, $[L^1]$ 收敛, 定理证毕.

**定理 3.2** 设 $X = \{X_n, \mathscr{F}_n, n = 0, 1, 2, \cdots, \infty\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅 (相应地, 上鞅), $\tau$ 和 $\eta$ 皆为 $\{\mathscr{F}_n\}$ 停时, $\tau \leqslant \eta$, 则 $\boldsymbol{E}(|X_\tau|) < \infty, \boldsymbol{E}(|X_\eta|) < \infty$, 且

$$\boldsymbol{E}(X_\eta|\mathscr{F}_\tau) = X_\tau \quad (\text{相应地}, \leqslant X_\tau). \tag{3.6}$$

**证** 设 $X$ 为鞅, 令 $\tau_n=\tau\mathbf{1}_{\{\tau\leqslant n\}}+\infty\cdot\mathbf{1}_{\{\tau>n\}}(n\geqslant 0)$, 推证:

$$X_{\tau_n}=\boldsymbol{E}(X_\infty|\mathscr{F}_{\tau_n})\quad(n\geqslant 0). \tag{3.7}$$

显然 $X_{\tau_n}\in\mathscr{F}_{\tau_n}$, 任取 $A\in\mathscr{F}_{\tau_n}$, 有

$$\int_A X_{\tau_n}dP=\sum_{k=0}^{n}\int_{A\cap\{\tau_n=k\}}X_k dP+\int_{A\cap\{\tau_n=\infty\}}X_\infty dP,$$

而 $X$ 是鞅, $A\bigcap\{\tau_n=k\}\in\mathscr{F}_k$, 所以

$$\int_{A\cap\{\tau_n=k\}}X_k dP=\int_{A\cap\{\tau_n=k\}}X_\infty dP\quad(0\leqslant k\leqslant n),$$

从而

$$\int_A X_{\tau_n}dP=\int_A X_\infty dP,$$

即 (3.7) 式成立. 任取 $\Omega$ 上的集合系 $\mathscr{G}$ 及 $\Omega$ 的子集 $A$, 记 $A\bigcap\mathscr{G}=\{B:B=A\bigcap D,D\in\mathscr{G}\}$, 则

$$\{\tau=\tau_n\}\bigcap\mathscr{F}_\tau=\{\tau=\tau_n\}\bigcap\mathscr{F}_{\tau_n}. \tag{3.8}$$

(事实上, 任取 $A\in\mathscr{F}_\tau$, 则 $\{\tau=\tau_n\}\bigcap A\bigcap\{\tau_n\leqslant k\}=\{\tau=\tau_n\}\bigcap(A\bigcap\{\tau\leqslant k\})\in\{\tau=\tau_n\}\bigcap\mathscr{F}_k$, 即 (3.8) 式的左边含于右边, 仿之可证右边含于左边.) 若注意 $\tau_n\geqslant\tau$, 从而 $\mathscr{F}_{\tau_n}\supset\mathscr{F}_\tau,\{\tau=\tau_n\}\in\mathscr{F}_\tau\subset\mathscr{F}_{\tau_n}$, 由 (3.8) 式得 $\mathscr{F}\supset\{\tau=\tau_n\}\bigcap\mathscr{F}_\tau=\{\tau=\tau_n\}\bigcap\mathscr{F}_{\tau_n}$. 再用条件期望性质及 (3.7) 式得

$$\boldsymbol{E}(X_\infty|\mathscr{F}_\tau)\mathbf{1}_{\{\tau=\tau_n\}}=X_\tau\mathbf{1}_{\{\tau=\tau_n\}}. \tag{3.9}$$

由于 $\{\tau=\tau_n\}\uparrow\Omega$, 所以在 (3.9) 式中令 $n\to\infty$, 得

$$\boldsymbol{E}(X_\infty|\mathscr{F}_\tau)=X_\tau. \tag{3.10}$$

故 $\boldsymbol{E}(|X_\tau|)<\infty$. 仿可证 $\boldsymbol{E}(|X_\eta|)<\infty$,

$$\boldsymbol{E}(X_\infty|\mathscr{F}_\eta)=X_\eta. \tag{3.11}$$

所以

$$\boldsymbol{E}(X_\eta|\mathscr{F}_\tau)=\boldsymbol{E}(\boldsymbol{E}(X_\infty|\mathscr{F}_\eta)|\mathscr{F}_\tau)=\boldsymbol{E}(X_\infty|\mathscr{F}_\tau)=X_\tau. \tag{3.12}$$

设 $X$ 为上鞅. 令 $Y_n=\boldsymbol{E}(X_\infty|\mathscr{F}_n),Z_n=X_n-Y_n(n=0,1,\cdots),Y_\infty=X_\infty,Z_\infty=0$, 则 $\{Y_n,\mathscr{F}_n,n=0,1,\cdots,\infty\}$ 为鞅, $\{Z_n,\mathscr{F}_n,n=0,1,\cdots,\infty\}$ 为非

负上鞅, 仿 (1.11)′ 式可证 $\boldsymbol{E}(Z_{\tau_n}) \leqslant \boldsymbol{E}(Z_0)$. 故由 Fatou 引理可知 $\boldsymbol{E}(Z_\tau) < \infty$, 从而 $X_t = Y_\tau + Z_\tau$ 的期望存在, 仿之 $\boldsymbol{E}(|X_\eta|) < \infty$. 再令

$$\eta_n = \eta \mathbf{1}_{\{\eta \leqslant n\}} + \infty \mathbf{1}_{\{\eta > n\}},$$

仿 (3.7) 式可证

$$Z_{\tau_n} \geqslant \boldsymbol{E}(Z_{\eta_n} | \mathscr{F}_{\tau_n}) \quad (n \geqslant 0). \tag{3.13}$$

从而

$$Z_\tau \mathbf{1}_{\{\tau = \tau_n\}} \geqslant \boldsymbol{E}(Z_{\eta_n} | \mathscr{F}_{\tau_n}) \mathbf{1}_{\{\tau = \tau_n\}} = \boldsymbol{E}(Z_{\eta_n} | \mathscr{F}_\tau) \mathbf{1}_{\{\tau = \tau_n\}}.$$

在上式中令 $n \uparrow \infty$, 并注意 $Z_{\eta_n} \uparrow Z_\eta, \{\tau = \tau_n\} \uparrow \Omega$, 可得

$$Z_\tau \geqslant \boldsymbol{E}(Z_\eta | \mathscr{F}_\tau). \tag{3.14}$$

但以 $Y$ 取代 $X$, 且由 (3.12) 式有

$$\boldsymbol{E}(Y_\eta | \mathscr{F}) = Y_\tau. \tag{3.12$'$}$$

由 (3.14)、(3.12)′ 式得 $X_\tau \geqslant \boldsymbol{E}(X_\eta | \mathscr{F}_\tau)$. 定理证毕.

**定理 3.3** 设 $X = \{X_n, \mathscr{F}_n, n = 0, 1, 2, \cdots, \infty\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅 (相应地, 上鞅), $\tau_0, \tau_1 . \tau_2, \cdots, \tau_\infty$ 均为 $\{\mathscr{F}_n\}$ 停时, 且 $\tau_0 \leqslant \tau_1 \leqslant \tau_2 \leqslant \cdots \leqslant \tau_\infty, \tau_i : \Omega \mapsto \{0, 1, 2, \cdots, \infty\} (i = 0, 1, 2 \cdots, \infty)$, 则

$$\{X_{\tau_n}, \mathscr{F}_{\tau_n}, n = 0, 1, \cdots, \infty\}$$

是鞅 (相应地, 上鞅).

**证** 由定理 3.2 有

$$\boldsymbol{E}(|X_{\tau_n}|) < \infty \quad (n = 0, 1, \cdots, \infty);$$

$\boldsymbol{E}(X_{\tau_{n+m}} | \mathscr{F}_{\tau_n}) = X_{\tau_n}$ (对应地 $\leqslant X_{\tau_n}$) $(n, m = 0, 1, \cdots, \infty)$, 故 $\{X_{\tau_n}, \mathscr{F}_{\tau_n}, n = 0, 1, 2, \cdots, \infty\}$ 是鞅 (相应地, 上鞅).

设 $(\Omega, \mathscr{F}, \mu)$ 为任一测度空间, $A, B \in \mathscr{F}$. 当 $\mu(A - B) = 0$ 时, 记 $A \subset B$, a.s. 当 $\mu(A \Delta B) = 0$ 时, 记 $A = B$, a.s., 有时记 "$A \subset B$, a.s." 为 "$A$ a.s. in $B$".

**定理 3.4** 设 $X = \{X_n, \mathscr{F}_n, n = 0, 1, 2, \cdots\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $D_n = X_n - X_{n-1} (n \geqslant 0, X_{-1} \equiv 0), D^* = \sup\limits_{n \geqslant 0} |D_n|, \boldsymbol{E}(D^*) < \infty$, 则下列三集合的概率相等

(i) $A_1 = \left\{\omega : \lim\limits_{n \to \infty} X_n(\omega)\ 收敛且有穷\right\}$;

(ii) $A_2 = \left\{\omega : \sup_{n\geqslant 0} X_n(\omega) < \infty\right\}$;

(iii) $A_3 = \left\{\omega : \inf_{n\geqslant 0} X_n(\omega) > -\infty\right\}$.

**证** 因为 "$\lim_{n\to\infty} X_n(\omega)$ 收敛且有穷 $\Leftrightarrow$ $\lim_{n\to\infty}(-X_n(\omega))$ 收敛且有穷", 所以 "$A_1 = A_2$, a.s. $\Leftrightarrow$ $A_1 = A_3$, a.s.". 又因为

$$A_1 \subset A_2,$$

所以为证定理 3.4, 只需证明 $A_2 \subset A_1$, a.s.

令 $\lambda > 0, \tau = \tau(\lambda) = \inf\{n : n \geqslant 0, X_n > \lambda\}$ (空集的 inf 定义为 $\infty$), 则由定理 3.1 得知 $X^{\tau(\lambda)}$ 是鞅. 若能证 $X^{\tau(\lambda)}$ 是 $L^1$ 有界的, 即

$$\|X^\tau\|_1 = \sup_{n\geqslant 0}\|X_n^\tau\|_1 = \sup_{n\geqslant 0}\|X_{n\wedge\tau}\|_1 < \infty,$$

则由定理 2.1 得知, 对 a.s. 的 $\omega$, $\lim_{n\to\infty} X_n^{\tau(\lambda)}(\omega)$ 收敛且有穷. 但是若令 $A_2(\lambda) = \left\{\omega : \sup_{n\geqslant 0} X_n(\omega) < \lambda\right\}$, 则由 $X_n^\tau = X_{n\wedge\tau}$ 得知

$$A_2(\lambda) \subset \{\tau(\lambda) = \infty\} \subset \bigcap_{n=0}^{\infty}\{X_n^{\tau(\lambda)} = X_n\},$$

所以

$$\begin{aligned}
A_2(\lambda) &\subset A_2(\lambda)\bigcap\left\{\lim_{n\to\infty} X_n^{\tau(\lambda)} \text{ 收敛且有穷}\right\}\\
&\subset \bigcap_{n=0}^{\infty}\{X_n^{\tau(\lambda)} = X_n\}\bigcap\left\{\lim_{n\to\infty} X_n^{\tau(\lambda)} \text{ 收敛且有穷}\right\}\\
&\subset A_1, \text{a.s.} \quad (\text{对一切 } \lambda > 0),
\end{aligned}$$

从而 $A_2 \subset A_1$, a.s..

下面补证 $X^{\tau(\lambda)}$ 是 $L^1$ 有界的. 因为

$$X_{n\wedge\tau} = X_n^\tau = \begin{cases} X_n \leqslant \lambda, & \text{当 } n < \tau,\\ X_\tau = X_{\tau-1} + D_\tau \leqslant \lambda + D^*, & \text{当 } n \geqslant \tau,\end{cases}$$

所以 $X_{n\wedge\tau} \leqslant \lambda + D^*$, 从而

$$\boldsymbol{E}(X_{n\wedge\tau}^+) \leqslant \boldsymbol{E}((\lambda + D^*) \vee 0) = \lambda + \boldsymbol{E}(D^*).$$

但是, 由定理 3.1 有

$$\boldsymbol{E}(X_0) = \boldsymbol{E}(X_{n\wedge\tau}) = \boldsymbol{E}(X_{n\wedge\tau}^+) - \boldsymbol{E}(X_{n\wedge\tau}^-),$$

所以

$$\begin{aligned}\boldsymbol{E}(|X_{n\wedge\tau}|) &= \boldsymbol{E}(X_{n\wedge\tau}^{+}) + \boldsymbol{E}(X_{n\wedge\tau}^{-}) = 2\boldsymbol{E}(X_{n\wedge\tau}^{+}) - \boldsymbol{E}(X_0) \\ &\leqslant 2(\lambda + \boldsymbol{E}(D^*)) - \boldsymbol{E}(X_0),\end{aligned}$$

从而 $\|X^\tau\|_1 = \sup\limits_{n\geqslant 0} \boldsymbol{E}(|X_{n\wedge\tau}|) < \infty$.

**命题 3.6**　设 $\xi$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上任一随机变量, $\boldsymbol{E}(|\xi|) < \infty$, 则 $\Omega$ 上的函数族

$$\{\boldsymbol{E}(\xi|\mathscr{G}) : \mathscr{G} \text{ 是 } \mathscr{F} \text{ 中任意子 } \sigma \text{ 代数}\}$$

是一致可积的.

**证**　不失普遍性可设 $\xi \geqslant 0$. 令 $F(\mathscr{G}) = \boldsymbol{E}(\xi|\mathscr{G})$, 则由 Chebyshev 不等式得

$$\begin{aligned}\int_{\{F(\mathscr{G})>K\}} F(\mathscr{G})dP &= \int_{\{F(\mathscr{G})>K\}} \xi dP \\ &= \int_{\{F(\mathscr{G})>K\}\cap\{\xi\leqslant J\}} \xi dP + \int_{\{F(\mathscr{G})>K\}\cap\{\xi>J\}} \xi dP \\ &\leqslant JP(F(\mathscr{G}) > K) + \int_{\{\xi>J\}} \xi dP \\ &\leqslant \frac{J}{K}\boldsymbol{E}(F(\mathscr{G})) + \int_{\{\xi>J\}} \xi dP \\ &= \frac{J}{K}\boldsymbol{E}(\xi) + \int_{\{\xi>J\}} \xi dP.\end{aligned}$$

因此

$$\lim_{K\to\infty} \sup_{\mathscr{G}\subset\mathscr{F}} \int_{\{F(\mathscr{G})>K\}} F(\mathscr{G})dP \leqslant \lim_{J\to\infty} \int_{\{\xi>J\}} \xi dP = 0,$$

即 $\{\boldsymbol{E}(\xi|\mathscr{G}) : \mathscr{G} \subset \mathscr{F}, \mathscr{G}$ 是 $\sigma$ 代数 $\}$ 是一致可积的.

**命题 3.7**　设 $\{X_n : n \geqslant 0\}, \{Y_n : n \geqslant 0\}$ 是概率空间 $(\Omega,\mathscr{F},P)$ 上两族随机变量, $\{Y_n : n \geqslant 0\}$ 是一致可积的, $\boldsymbol{E}(|X_n|) < \infty (n \geqslant 0)$.

(1) 若 $X_n \geqslant Y_n$, a.s. $(n \geqslant 0)$, 则

$$\boldsymbol{E}\left(\liminf_{n\to\infty} X_n\right) \leqslant \lim_{n\to\infty} \inf \boldsymbol{E}(X_n).$$

(2) 若 $X_n \leqslant Y_n$, a.s. $(n \geqslant 0)$, 则

$$\limsup_{n\to\infty} \boldsymbol{E}(X_n) \leqslant \boldsymbol{E}\left(\limsup_{n\to\infty} X_n\right).$$

**证** 令 $X_n^a = X_n \vee a$ ($a$ 是实数), 则由 Fatou 引理有

$$\boldsymbol{E}\left(\liminf_{n\to\infty} X_n\right) \leqslant \boldsymbol{E}\left(\liminf_{n\to\infty} X_n^a\right) \leqslant \liminf_{n\to\infty} \boldsymbol{E}(X_n^a). \tag{3.15}$$

但是

$$\boldsymbol{E}(X_n^a - X_n) = \int_{\{X_n<a\}} (a - X_n)dP,$$

且 $a<0$ 时, 由 $X_n \geqslant Y_n$, 有

$$\begin{aligned}\left|\int_{\{X_n<a\}} adP\right| &\leqslant \int_{\{X_n<a\}} |X_n|dP \leqslant \int_{\{Y_n<a\}} |Y_n|dP \\ &\leqslant \int_{\{|Y_n|>|a|\}} |Y_n|dP;\end{aligned}$$

$$\left|\int_{\{X_n<a\}} X_n dP\right| \leqslant \int_{\{X_n<a\}} |X_n|dP \leqslant \int_{\{|Y_n|>|a|\}} |Y_n|dP.$$

而 $\{Y_n : n \geqslant 0\}$ 是一致可积的, 所以对任何 $\varepsilon > 0$, 存在 $A > 0$, 使 $|a| > A$ 时, 有

$$\int_{\{|Y_n|>|a|\}} |Y_n|dP < \frac{\varepsilon}{2} \quad (\text{对一切 } n \geqslant 0).$$

所以 “$a < -A \Rightarrow |\boldsymbol{E}(X_n^a - X_n)| < \varepsilon$ (对一切 $n \geqslant 0$)”.

$$\liminf_{n\to\infty} \boldsymbol{E}(X_n^a) \leqslant \liminf_{n\to\infty} \boldsymbol{E}(X_n) + \varepsilon \quad (a < -A), \tag{3.16}$$

由 $\varepsilon < 0$ 可任意小及 (3.15)、(3.16) 式得

$$\boldsymbol{E}\left(\liminf_{n\to\infty} X_n\right) \leqslant \liminf_{n\to\infty} \boldsymbol{E}(X_n).$$

仿之可证

$$\limsup_{n\to\infty} \boldsymbol{E}(X_n) \leqslant \boldsymbol{E}\left(\limsup_{n\to\infty} X_n\right).$$

**命题 3.8** 设 $X = \{X_n, \mathscr{F}_n, n = 0, -1, -2, \cdots\}$ 为概率空间 $(\Omega, \mathscr{F}, P)$ 上的上鞅. 如果 $\lim\limits_{n\to\infty} \boldsymbol{E}(X_n) = \sup\limits_{n\leqslant 0} \boldsymbol{E}(X_n) = A < \infty$, 则 $\{X_n : n = 0, -1, -2, \cdots\}$ 是致可积的.

**证** 设 $X$ 为上鞅, 则 $\{Y_n = X_n - \boldsymbol{E}(X_0|\mathscr{F}_n), \mathscr{F}_n, n = 0, -1, -2, \cdots\}$ 为非负上鞅, 而由命题 3.6 得知 $\{\boldsymbol{E}(X_0|\mathscr{F}_n) : n = 0, -1, -2, \cdots\}$ 是一致可积的, 所以为证 $\{X_n : n = 0, -1, -2, \cdots\}$ 一致可积, 只需证 $\{Y_n : n = 0, -1, -2, \cdots\}$ 一致可积, 由

$$\lim_{n\to\infty} \boldsymbol{E}(X_n) = \sup_{n\leqslant 0} \boldsymbol{E}(X_n) = A < \infty,$$

$$\lim_{n\to\infty} \boldsymbol{E}(Y_n) = \sup_{n\leqslant 0} \boldsymbol{E}(Y_n) = A - \boldsymbol{E}(X_0) \equiv B < \infty.$$

所以任给 $\varepsilon > 0$, 可取自然数 $K$, 使

$$B - \boldsymbol{E}(Y_{-K}) < \frac{\varepsilon}{2}. \tag{3.17}$$

对 $c > 0$ 及 $n < -K$, 由上鞅性得

$$\begin{aligned}\int_{\{Y_n > c\}} Y_n dP &= \boldsymbol{E}(Y_n) - \int_{\{Y_n \leqslant c\}} Y_n dP \\ &\leqslant \boldsymbol{E}(Y_n) - \int_{\{Y_n \leqslant c\}} Y_{-K} dP \\ &= \boldsymbol{E}(Y_n) - \boldsymbol{E}(Y_{-K}) + \int_{\{Y_n > c\}} Y_{-K} dP,\end{aligned} \tag{3.18}$$

由于 $B \geqslant \boldsymbol{E}(Y_n) \geqslant \boldsymbol{E}(Y_{-K}) (n < -K)$, 所以由 (3.17) 式得

$$\boldsymbol{E}(Y_n) - \boldsymbol{E}(Y_{-K}) < \frac{\varepsilon}{2} \quad (n < -K). \tag{3.19}$$

又因为

$$P(Y_n > c) \leqslant \frac{1}{c} \boldsymbol{E}(Y_n) \leqslant \frac{B}{c} \quad (n \leqslant 0), \tag{3.20}$$

从而当 $c$ 充分大时, 有

$$\int_{\{Y_n > c\}} Y_{-K} dP < \frac{\varepsilon}{2} \quad (n \geqslant 0). \tag{3.21}$$

由 (3.18), (3.19), (3.21) 式得

$$\sup_{n < -K} \int_{\{Y_n > c\}} Y_n dP < \varepsilon \quad (c \text{ 充分大})$$

而 $0, -1, -2, \cdots, -K$ 只有有限个数, 所以, 可取 $c$ 充分大, 使

$$\sup_{n \geqslant 0} \int_{\{Y_n > c\}} Y_n dP < \varepsilon.$$

故 $\{Y_n : n = 0, -1, -2, \cdots\}$ 是一致可积的.

下面我们研究连续时间参数的鞅的 Doob 停时理论.

**定理 3.5** 设 $X = \{X_t, \mathscr{F}_t, t \in [0, \infty]\}$ 为概率空间 $(\Omega, \mathscr{F}, P)$ 上的上鞅 (鞅), 其所有的轨道 $X(\bullet, \omega)$ 右连续, $\tau, \eta$ 皆为 $\{\mathscr{F}_t\}$ 停时, $\tau \leqslant \eta$, 则 $\boldsymbol{E}(|X_\tau|) < \infty, \boldsymbol{E}(|X_\eta|) < \infty$, 且

$$\boldsymbol{E}(X_\eta | \mathscr{F}_\tau) \leqslant \boldsymbol{E}(X_\tau | \mathscr{F}_\tau) = X_\tau, \tag{3.22}$$

$$(\boldsymbol{E}(X_\eta | \mathscr{F}_\tau) = \boldsymbol{E}(X_\tau | \mathscr{F}_\tau) = X_\tau).$$

**证** 由于 $X(\bullet,\omega)$ 右连续, 所以 $\{X_t, t\in[0,\infty]\}$ 关于 $\{\mathscr{F}_t, t\in[0,\infty]\}$ 是循序可测的. 因此, (参见 [35] 第七章命题 3.6) $X_\tau\in\mathscr{F}_\tau$, 从而 $\boldsymbol{E}(X_\tau|\mathscr{F}_\tau)=X_\tau$.

因 $X$ 是上鞅, 推证 (3.22) 式, 令

$$\boldsymbol{T}_n=\left\{0,\frac{1}{2^n},\frac{2}{2^n},\cdots,\infty\right\}\quad(n\geqslant 0),$$

则 $\{X_t,\mathscr{F}_t,t\in\boldsymbol{T}_n\}$ 是上鞅, 再令

$$\tau_n=\sum_{k=1}^{\infty}\frac{k}{2^n}\mathbf{1}\left\{\frac{k-1}{2^n}\leqslant\tau<\frac{k}{2^n}\right\}+\infty\bullet\mathbf{1}_{\{\tau=\infty\}},\tag{3.23}$$

$\eta_n$ 之定义仿 (3.23) 式, 则 $\tau_n,\eta_n$ 皆为 $\{\mathscr{F}_t:t\in\boldsymbol{T}_n\}$ 停时, 且 $\tau_n\downarrow\tau,\eta_n\downarrow\eta$. 所以由定理 3.2 得知

$$\boldsymbol{E}(X_{\eta_n}|\mathscr{F}_{\tau_n})\leqslant X_{\tau_n};\tag{3.24}$$

$$\boldsymbol{E}(|X_{\tau_n}|)<\infty,\quad \boldsymbol{E}(|X_{\eta_n}|)<\infty\quad(n\geqslant 0).$$

由命题 3.4 有 $\mathscr{F}_\tau\subset\mathscr{F}_{\tau_n}(n\geqslant 0)$, 所以由 (3.24) 式得知

$$\int_A X_{\eta_n}dP\leqslant\int_A X_{\tau_n}dP\quad(A\in\mathscr{F}_\tau\subset\mathscr{F}_{\tau_n}).\tag{3.25}$$

而由 $X(\bullet,\omega)$ 右连续及 $\tau_n\downarrow\tau,\eta_n\downarrow\eta$ 得

$$\lim_{n\to\infty}X_{\eta_n}=X_\eta,\quad \lim_{n\to\infty}X_{\tau_n}=X_\tau,\tag{3.26}$$

若能证 $\{X_{\eta_n}:n=0,1,2,\cdots\},\{X_{\tau_n}:n=0,1,2,\cdots\}$, 皆一致可积, 则由引理 1.3 得知

$$\lim_{n\to\infty}X_{\eta_n}=X_\eta,L_1;\quad \lim_{n\to\infty}X_{\tau_n}=X_\tau,L^1.$$

从而有

$$\lim_{n\to\infty}\mathbf{1}_AX_{\eta_n}=\mathbf{1}_AX_\eta,\quad L^1.(A\in\mathscr{F}_\tau);$$

$$\lim_{n\to\infty}\mathbf{1}_AX_{\tau_n}=\mathbf{1}_AX_\tau,\quad L^1.(A\in\mathscr{F}_\tau).$$

且有

$$\begin{aligned}\int_A X_\eta dP&=\boldsymbol{E}(\mathbf{1}_AX_\eta)=\lim_{n\to\infty}\boldsymbol{E}(\mathbf{1}_AX_{\eta_n})\\&\leqslant\lim_{n\to\infty}\boldsymbol{E}(\mathbf{1}_AX_{\tau_n})=\boldsymbol{E}(\mathbf{1}_AX_\tau)\\&=\int_A X_\tau dP\quad(A\in\mathscr{F}_\tau)\end{aligned}\tag{3.27}$$

由 $X_\tau \in \mathscr{F}_\tau$ 及 (3.27) 式即得 (3.22) 式.

下面补证 $\{X_{\tau_n}: n=0,1,2,\cdots\}$ 一致可积 ($\{X_{\eta_n}: n=0,1,2,\cdots\}$ 类似). 由于 $\tau_n$ 与 $\tau_{n+1}$ 皆在 $\boldsymbol{T}_{n+1}$ 中取值, 且 $\tau_n \geqslant \tau_{n+1}$, 所以由定理 3.2 得

$$\boldsymbol{E}(X_{\tau_n}|\mathscr{F}_{\tau_{n+1}}) \leqslant X_{\tau_{n+1}} \quad (n \geqslant 0).$$

令

$$Y_{-n} = X_{\tau_n}, \quad \mathscr{G}_{-n} = \mathscr{F}_{\tau_n} \quad (n \geqslant 0),$$

则 $\{Y_{-n}, \mathscr{G}_{-n}, n=0,1,2,\cdots\}$ 是上鞅. 又因为 $\boldsymbol{E}(Y_{-n}) = \boldsymbol{E}(X_{\tau_n}) \leqslant \boldsymbol{E}(X_0)(n \geqslant 0)$, 所以由命题 3.8 得知 $\{Y_{-n} = X_{\tau_n}: n=0,1,2,\cdots\}$ 是一致可积的, 定理证毕.

**定理 3.6** 设 $X=\{X_t, \mathscr{F}_t, t\in[0,\infty]\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的上鞅 (鞅), 其所有的轨道 $X(\cdot,\omega)$ 右连续, $\tau_s$ 皆为 $\{\mathscr{F}_t\}$ 停时, 且当 $s_1<s_2, s_1,s_2\in[0,\infty]$ 时, 有 $\tau_{s_2} \geqslant \tau_{s_1}$, 则 $\{X_{\tau_t}, \mathscr{F}_{\tau_t}, t\in[0,\infty]\}$ 是上鞅 (鞅).

**证** 由定理 3.5 即得定理 3.6.

**定理 3.7** 设 $X=\{X_t, \mathscr{F}_t, t\in[0,\infty]\}$ 为概率空间 $(\Omega,\mathscr{F},P)$ 上的鞅 (上鞅), 其所有的轨道 $X(\cdot,\omega)$ 右连续, 且 $\{X_t: t\in[0,\infty)\}$ 一致可积, $\tau$ 是 $\{\mathscr{F}_{t+}\}$ 停时, 则 $X^\tau = \{X_{t\wedge\tau}, \mathscr{F}_t, t\in[0,\infty)\}$ 是鞅 (上鞅).

**证** 设 $X$ 是鞅. 令 $\boldsymbol{T}_n = \left\{0, \dfrac{1}{2^n}, \dfrac{2}{2^n}, \cdots, \infty\right\}$, $\tau_n$ 如 (3.23) 式所定义 ($n \geqslant 0$), 则 $\tau_n \downarrow \tau$, 且 $\tau_n: \Omega \mapsto \boldsymbol{T}_n$,

$$\left\{\tau_n = \frac{k}{2^n}\right\} = \left\{\frac{k-1}{2^n} \leqslant \tau < \frac{k}{2^n}\right\} \in \mathscr{F}_{k/2^n},$$

所以 $\tau_n$ 是 $\{\mathscr{F}_t: t\in\boldsymbol{T}_n\}$ 停时. 显然 $\tau_n$ 亦可视为 $\{\mathscr{F}_t: t\in[0,\infty)\}$ 停时, 再用 $\tau_n\downarrow\tau, X(\cdot,\omega)$ 右连续知 $\lim\limits_{n\to\infty} X_{\tau_n} = X_\tau$. 所以

$$\begin{aligned}
\{X_{\tau\wedge t}<\lambda\} &= \{X_t<\lambda, \tau\geqslant t\} \bigcup \{X_\tau<\lambda, \tau<t\} \\
&= \{X_t<\lambda, \tau\geqslant t\} \bigcup \left(\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{j=n}^{\infty}\left\{X_{\tau_j}<\lambda-\frac{1}{m}, \tau_j<t\right\}\right) \\
&= \{X_t<\lambda, \tau\geqslant t\} \bigcup \left(\bigcup_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcup_{j=n}^{\infty}\bigcap_{\frac{k}{2^j}<t}\left\{X_{k/2^j}<\lambda-\frac{1}{m}, \tau_j=\frac{k}{2^j}\right\}\right) \\
&\in \mathscr{F}_t \quad (t\in[0,\infty)).
\end{aligned}$$

由 $\{X_t: t\in[0,\infty)\}$ 是一致可积鞅、所有轨道右连续及定理 2.2′ 得知: 存在随机变量 $X_\infty, \boldsymbol{E}(|X_\infty|)<\infty$, 使

$$\lim_{t\to\infty} X_t = X_\infty, \quad \text{a.s. 和 } L^1., \tag{3.28}$$

且 $\{X_t, \mathscr{F}_t, t \in [0, \infty]\}$ 是鞅, 在定理 3.5 中取 $\eta \equiv \infty, \tau = \tau_n \wedge t$, 得

$$E(X_\infty | \mathscr{F}_{\tau_n \wedge t}) = X_{\tau_n \wedge t}.$$

再用命题 3.6 得知 $\{X_{\tau_n \wedge t} : t \in [0, \infty], n \geqslant 0\}$ 是一致可积的. 由 $\lim\limits_{n \to \infty} X_{\tau_n \wedge t} = X_{\tau \wedge t}$ 及引理 1.3 得知

$$\lim_{n \to \infty} X_{\tau_n \wedge t} = X_{\tau \wedge t}, \quad L^1., \tag{3.29}$$

且有

$$\lim_{n \to \infty} \int_A X_{\tau_n \wedge t} dP = \int_A X_{\tau \wedge t} dP (t \in [0, \infty], A \in \mathscr{F}). \tag{3.30}$$

由 (3.29) 及 (3.30) 式, 为证 $X^\tau$ 是鞅, 只需证明

$$\int_A X_{\tau_n \wedge s} dP = \int_A X_{\tau_n \wedge t} dP \quad (A \in \mathscr{F}_s, s < t, n \geqslant 0). \tag{3.31}$$

今任取 $A \in \mathscr{F}_s, s < t, n \geqslant 0$, 由 $X$ 是鞅得

$$\begin{aligned}
&\int_A X_{\tau_n \wedge s} dP \\
&= \int_{A \cap \{\tau_n < s\}} X_{\tau_n} dP + \int_{A \cap \{\tau_n \geqslant s\}} X_s dP \\
&= \int_{A \cap \{\tau_n < s\}} X_{\tau_n} dP + \int_{A \cap \{\tau_n \geqslant s\}} X_t dP \\
&= \int_{A \cap \{\tau_n < s\}} X_{\tau_n} dP + \sum_{t > \frac{k}{2^n} \geqslant s} \int_{A \cap \left\{\tau_n = \frac{k}{2^n}\right\}} X_t dP + \sum_{\frac{k}{2^n} \geqslant t} \int_{A \cap \left\{\tau_n = \frac{k}{2^n}\right\}} X_t dP \\
&= \int_{A \cap \{\tau_n < t\}} X_{\tau_n} dP + \sum_{t > \frac{k}{2^n} \geqslant s} \int_{A \cap \left\{\tau_n = \frac{k}{2^n}\right\}} (X_t - X_{\tau_n}) dP + \int_{A \cap \{\tau_n \geqslant t\}} X_t dP \\
&= \int_{A \{\tau_n \geqslant t\}} X_t dP + \sum_{t > \frac{k}{2^n} \geqslant s} \int_{A \cap \left\{\tau_n = \frac{k}{2^n}\right\}} (X_t - X_{k/2^n}) dP
\end{aligned} \tag{3.32}$$

由 $X$ 是鞅,$A \in \mathscr{F}_s \subset \mathscr{F}_{k/2^n} \left( s \leqslant \dfrac{k}{2^n} \text{ 时} \right), \left\{ \tau_n = \dfrac{k}{2^n} \right\} \in \mathscr{F}_{k/2^n}$, 可知 $t > \dfrac{k}{2^n} \geqslant s$ 时, $A \bigcap \left\{ \tau_n = \dfrac{k}{2^n} \right\} \in \mathscr{F}_{k/2^n}$, 从而

$$\int_{A \cap \left\{\tau_n = \frac{k}{2^n}\right\}} (X_t - X_{k/2^n}) dP = 0. \tag{3.33}$$

由 (3.32)、(3.33) 得 (3.31) 式. 定理证毕.

# *§4 鞅 变 换

上世纪 60 年代, 由于调和分析及 Banach 空间中的几何学对鞅的需要, 以及它们相互渗透所产生的新课题, 鞅变换得到了快速发展. **鞅变换**即是鞅与某种随机变量列的组合.

**定义 4.1** 设 $(\Omega, \mathscr{F}, P)$ 为概率空间, $\{\mathscr{F}_n : n = -1, 0, 1, 2, \cdots\}$ 为 $\mathscr{F}$ 中的一个非降子 $\sigma$ 代数族, 若

$$V_n : \Omega \mapsto \mathbf{R}, \quad V_n \in \mathscr{F}_{n-1} \quad (n = 0, 1, \cdots), \tag{4.1}$$

则称 $V = \{V_n : n \geqslant 0\}$ 为 $\{\mathscr{F}_n\}$ **可预报序列**; 若 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 为下鞅 (或上鞅), $D_n = X_n - X_{n-1} (n \geqslant 0, X_{-1} \equiv 0), V$ 为 $\{\mathscr{F}_n\}$ 可预报序列, 则称

$$Y = \{Y_n, \mathscr{F}_n, n \geqslant 0\} \left( Y_n = \sum_{k=0}^{n} V_k D_k, n \geqslant 0 \right) \tag{4.2}$$

为 $X$ 的关于**可预报序列 $V$ 的变换**.

**例 4.1** 设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 为鞅, $\tau$ 为 $\{\mathscr{F}_n\}$ 停时, $\tau : \Omega \mapsto \{0, 1, 2, \cdots, \infty\}$. 令

$$V_n = \mathbf{1}_{\{\tau \geqslant n\}} \quad (n \geqslant 0),$$

$$\mathscr{F}_{-1} = \{\varnothing, \Omega\}, \quad D_n = X_n - X_{n-1} \quad (n \geqslant 0, X_{-1} \equiv 0),$$

则 $V = \{V_n : n \geqslant 0\}$ 为 $\{\mathscr{F}_n\}$ 可预报序列,

$$Y_n = \sum_{k=0}^{n} V_k D_k = \sum_{k=0}^{n} \mathbf{1}_{\{\tau \geqslant k\}} D_k = X_{n \wedge \tau},$$

所以 $X$ 关于 $V$ 的变换 $Y = X^{\tau}, Y$ 是一个鞅.

注意: 鞅 $X$ 关于可预报序列 $V$ 的变换未必为鞅, 因为 $V_k D_k$ 未必可积.

但是, 若对每个 $n \geqslant 0, V_n$ 有界, 则鞅 $X$ 关于可预报序列 $V$ 的变换 $Y$ 必为鞅. 因为 $V_k D_k \in L^1(\Omega, \mathscr{F}, P), V_k D_k \in \mathscr{F}_k$, 所以 $Y_n = \sum\limits_{k=0}^{n} V_k D_k \in \mathscr{F}_n, Y_n \in L^1(\Omega, \mathscr{F}, P)$. 又因为

$$\boldsymbol{E}(V_{n+1} D_{n+1} | \mathscr{F}_n) = V_{n+1} \boldsymbol{E}(D_{n+1} | \mathscr{F}_n) = 0,$$

所以

$$\boldsymbol{E}(Y_{n+1} | \mathscr{F}_n) = \boldsymbol{E}(V_{n+1} D_{n+1} | \mathscr{F}_n) + \boldsymbol{E}(Y_n | \mathscr{F}_n) = Y_n.$$

即 $Y = \{Y_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅.

对每个 $n \geqslant 0, V_n$ 有界 (甚至 $\{V_n : n \geqslant 0\}$ 一致有界 1), $Y$ 的性质也比 $X$ 的性质要差. 下面的例子说明这一论断.

**例 4.2**　设 $\{X_n : n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的以

$$\{-b, -b+1, \cdots, -1, \cdots, b-1, b\}$$

为状态空间的时齐的 Markov 过程, 且 $|i| < b$ 时,

$$p_{i,j} = P(X_{n+1} = j | X_n = i) = \begin{cases} \dfrac{1}{2}, & \text{当 } j = i+1, \\ \dfrac{1}{2}, & \text{当 } j = i-1, \end{cases}$$

$|i| = b$ 时,

$$p_{i,i} = P(X_{n+1} = i | X_n = i) = 1, \quad P(X_0 = 0) = 1.$$

(这样的 Markov 过程称为在 $\pm b$ 具有吸收屏的简单随机徘徊). 令 $\mathscr{F}_n = \sigma(X_0, X_1, \cdots, X_n)(n \geqslant 0), \mathscr{F}_{-1} = \{\varnothing, \Omega\}$, 令 $V_k \equiv (-1)^{k+1} (k \geqslant 0)$, 则 $V = \{V_k : k \geqslant 0\}$ 是 $\{\mathscr{F}_n\}$ 可预报序列, $V$ 一致有界 1, $\{X_n : n \geqslant 0\}$ 一致有界 $b$. 显然

$$\boldsymbol{E}(X_{n+1} | \mathscr{F}_n) = \boldsymbol{E}(X_{n+1} | X_n) = X_n \quad (n \geqslant 0),$$

即 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅. 由 $|X_n| \leqslant b$ 得知

$$\|X\|_\infty = \sup_{n \geqslant 0} \|X_n\|_\infty \leqslant b < \infty.$$

若令

$$\Omega_n = \left\{\omega : X_0(\omega) = 0, X_1(\omega) = 1, \cdots, X_n(\omega) = \frac{1 + (-1)^{n+1}}{2}\right\},$$

则当 $\omega \in \Omega_n$ 时,

$$\begin{aligned} Y_n(\omega) &= \sum_{k=1}^{n} (-1)^{k+1} \left[\frac{1 + (-1)^{k+1}}{2} - \frac{1 + (-1)^k}{2}\right] + (-1)^1 D_0(\omega) \\ &= \sum_{k=1}^{n} \left(\frac{1}{2} + \frac{1}{2}\right) = n, \end{aligned}$$

所以 $P(Y_n = n) \geqslant P(\Omega_n) = \left(\frac{1}{2}\right)^n$, 因此

$$\begin{aligned}\|Y\|_\infty &= \sup_{n\geqslant 0}\|Y_n\|_\infty = \sup_{n\geqslant 0}\left(\lim_{p\to\infty}\left(\int_\Omega |Y_n|^p dP\right)^{\frac{1}{p}}\right)\\&\geqslant \sup_{n\geqslant 0}\left(\lim_{p\to\infty}\left(\int_{\Omega_n} |Y_n|^p dP\right)^{\frac{1}{p}}\right)\\&= \sup_{n\geqslant 0}\left(\lim_{p\to\infty}\left(\left(\frac{1}{2}\right)^n n^p\right)^{\frac{1}{p}}\right) = \infty.\end{aligned}$$

**定理 4.1 (Burkholder)** 设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $V = \{V_n : n \geqslant 0\}$ 是 $\{\mathscr{F}_n\}$ 可预报序列, $D_n = X_n - X_{n-1} (n \geqslant 0, X_{-1} \equiv 0), Y = \{Y_n, \mathscr{F}_n, n \geqslant 0\}$ 是 $X$ 关于 $V$ 的变换, 即

$$Y_n = \sum_{k=0}^n V_k D_k \quad (n \geqslant 0), \tag{4.3}$$

则 "$\|X\|_1 < \infty \Rightarrow \lim\limits_{n\to\infty} Y_n$ 在 $\{V^* < \infty\}$ 上是 a.s. 收敛且有穷", 其中 $V^* = \sup\limits_{n\geqslant 0}|V_n|$.

特别地,

$$\text{“}\|X\|_1 < \infty, \quad V^* \leqslant 1 \Rightarrow \lim_{n\to\infty} Y_n, \quad \text{a.s. 收敛且有穷”},$$

**证** 我们分几步来证明此定理.

(I) 设 $\|X\|_2 < \infty, V^* \leqslant 1$. 推证: $\lim\limits_{n\to\infty} Y_n$, a.s. 收敛且有穷. 由于 $V_k \in \mathscr{F}_{k-1} (k \geqslant 0), D_k \in \mathscr{F}_k (k \geqslant 0)$,

$$\boldsymbol{E}(D_k|\mathscr{F}_{k-1}) = 0 \quad (k \geqslant 1),$$

所以当 $j < k$ 时,

$$\begin{aligned}\boldsymbol{E}(V_j D_j V_k D_k) &= \boldsymbol{E}(\boldsymbol{E}(V_j D_j V_k D_k|\mathscr{F}_{k-1}))\\&= \boldsymbol{E}(V_j D_j V_k \boldsymbol{E}(D_k|\mathscr{F}_{k-1})) = 0.\end{aligned}$$

因此, 由上式及 $V^* \leqslant 1$ 和定理 1.1 得

$$\begin{aligned}\|Y_n\|_2^2 &= \left\|\sum_{k=0}^n V_k D_k\right\|_2^2 = \sum_{k=0}^n \|V_k D_k\|_2^2\\&\leqslant \sum_{k=0}^n \|D_k\|_2^2 = \|X_n\|_2^2.\end{aligned}$$

所以, 由 $\|X\|_2<\infty$ 得 $\|Y\|_2<\infty$, 更有 $\|Y\|_1<\infty$. 由 $V^*\leqslant 1$ 知 $Y$ 是鞅, 所以由定理 2.1 得知: $\lim\limits_{n\to\infty} Y_n$, a.s. 收敛且有穷.

(Ⅱ) 设 $X$ 是一致有界 $M$ (即 $\sup\limits_{n\geqslant 0}|X_n|\leqslant M$) 的下鞅, $V^*\leqslant 1$, 推证 $\lim\limits_{n\to\infty} Y_n$, a.s. 收敛且有穷. 令

$$\widehat{D}_0=D_0,\quad \widehat{D}_k=-\boldsymbol{E}(D_k|\mathscr{F}_{k-1})+D_k(k\geqslant 1),$$

则 $\boldsymbol{E}(\widehat{D}_k|\mathscr{F}_{k-1})=0(k\geqslant 1)$, 因此, 由定理 1.1 得知 $\{\widehat{D}_k:k\geqslant 0\}$ 是某个鞅 $\widehat{X}=\{\widehat{X}_n,\mathscr{F}_n,n\geqslant 0\}$ 的差序列. 令 $\widehat{Y}_n=\sum\limits_{k=0}^{n}V_k\widehat{D}_k(n\geqslant 0)$. 因为

$$\begin{aligned}\boldsymbol{E}(D_k\boldsymbol{E}(D_k|\mathscr{F}_{k-1}))&=\boldsymbol{E}(\boldsymbol{E}(D_k\boldsymbol{E}(D_k|\mathscr{F}_{k-1})|\mathscr{F}_{k-1}))\\&=\boldsymbol{E}(\boldsymbol{E}(D_k|\mathscr{F}_{k-1})^2),\end{aligned}\tag{4.4}$$

因此

$$\begin{aligned}\|\widehat{D}_k\|_2^2&=\|D_k-\boldsymbol{E}(D_k|\mathscr{F}_{k-1})\|_2^2\\&=\|D_k\|_2^2-2\boldsymbol{E}(D_k\boldsymbol{E}(D_k|\mathscr{F}_{k-1}))+\boldsymbol{E}(\boldsymbol{E}(D_k|\mathscr{F}_{k-1})^2)\\&=\|D_k\|_2^2-\boldsymbol{E}(\boldsymbol{E}(D_k|\mathscr{F}_{k-1})^2)\leqslant\|D_k\|_2^2.\end{aligned}\tag{4.5}$$

不妨令 $X$ 非负, 否则考虑 $X^*=\{X_n^*=X_n+M,\mathscr{F}_n,n\geqslant 0\}$, 则 $X^*$ 是非负的一致有界的下鞅, 而且 $X$ 与 $X^*$ 的差序列一样: $D_n=X_n-X_{n-1}=X_n^*-X_{n-1}^*=D_n^*$, 从而 $X$ 与 $X^*$ 关于 $V$ 的变换亦一样. 下面我们就假定 $X$ 是非负一致有界下鞅. 由下鞅性, 有

$$\boldsymbol{E}(D_n|\mathscr{F}_{n-1})=\boldsymbol{E}(X_n|\mathscr{F}_{n-1})-X_{n-1}\geqslant 0\quad(n\geqslant 1)\tag{4.6}$$

再用 $X_n\geqslant 0,(n\geqslant 0)$, 得

$$\begin{aligned}\boldsymbol{E}(|X_n|^2)&=\boldsymbol{E}(|X_{n-1}+D_n|^2)\\&=\boldsymbol{E}(|D_n|^2)+2\boldsymbol{E}(X_{n-1}D_n)+\boldsymbol{E}(|X_{n-1}|^2)\\&=\boldsymbol{E}(|D_n|^2)+2\boldsymbol{E}(X_{n-1}\boldsymbol{E}(D_n|\mathscr{F}_{n-1}))+\boldsymbol{E}(|X_{n-1}|^2)\\&\geqslant\boldsymbol{E}(|D_n|^2)+\boldsymbol{E}(|X_{n-1}|^2)\quad(n\geqslant 1),\end{aligned}$$

从而

$$\begin{aligned}\|X_n\|_2^2=\boldsymbol{E}(|X_n|^2)&=\sum_{k=0}^{n}(\boldsymbol{E}(|X_k|^2)-\boldsymbol{E}(|X_{k-1}|^2))\\&\geqslant\sum_{k=0}^{n}\boldsymbol{E}(|D_k|^2)=\sum_{k=0}^{n}\|D_k\|_2^2.\end{aligned}\tag{4.7}$$

而由定理 1.1 及 (4.5)、(4.7) 式有

$$\begin{aligned}\|\widehat{X}_n\|_2^2 = \sum_{k=0}^{n} \|\widehat{D}_k\|_2^2 \leqslant \sum_{k=0}^{n} \|D_k\|_2^2 \\ \leqslant \|X_n\|_2^2 \leqslant M^2 < \infty \quad (n \geqslant 0),\end{aligned} \tag{4.8}$$

所以 $\widehat{X}$ 是 $L^2$ 有界的, 因此由定理 2.1 及 (1), 得知

$$\lim_{n\to\infty} \widehat{X}_n \text{ 和 } \lim_{n\to\infty} \widehat{Y}_n$$

皆是 a.s. 收敛且有穷. 而由 (4.6) 式知

$$\boldsymbol{E}(D_k|\mathscr{F}_{k-1}) \geqslant 0 \quad (k \geqslant 1),$$

所以

$$\sum_{k=1}^{\infty} \boldsymbol{E}(D_k|\mathscr{F}_{k-1}) = Z \geqslant 0$$

收敛. 推证 $Z < \infty$, [a.e.]. 事实上, 由积分的单调收敛定理, 有

$$\boldsymbol{E}(|Z|^2) = \lim_{n\to\infty} \boldsymbol{E}\left(\left|\sum_{k=1}^{n} \boldsymbol{E}(D_k|\mathscr{F}_{k-1})\right|^2\right).$$

而由 $\widehat{D}_n$ 的定义有

$$\sum_{k=1}^{n} \boldsymbol{E}(D_k|\mathscr{F}_{k-1}) = \sum_{k=1}^{n}(D_k - \widehat{D}_k) = X_n - \widehat{X}_n \quad (n \geqslant 1),$$

所以

$$\left\|\sum_{k=1}^{n} \boldsymbol{E}(D_k|\mathscr{F}_{k-1})\right\|_2 \leqslant \|X_n\|_2 + \|\widehat{X}_n\|_2 \leqslant 2M \quad (n \geqslant 1),$$

从而 $\|Z\|_2 \leqslant 2M$, 故 $Z < \infty$, a.s., 又因为 $V^* \leqslant 1$, 所以

$$\sum_{k=1}^{\infty} |V_k \boldsymbol{E}(D_k|\mathscr{F}_{k-1})| < \infty, \quad \text{a.s..}$$

但是

$$Y_n = \widehat{Y}_n + \sum_{k=1}^{n} V_k \boldsymbol{E}(D_k|\mathscr{F}_k) \quad (n \geqslant 1),$$

所以由 $\lim\limits_{n\to\infty} \widehat{Y}_n$ 是 a.s. 收敛且有穷得知 $\lim\limits_{n\to\infty} Y_n$, a.s. 收敛且有穷.

(Ⅲ) 设 $X$ 是非负 $L^1$ 有界鞅, $V^* \leqslant 1$, 推证 $\lim\limits_{n\to\infty} Y_n$, a.s. 收敛且有穷.

任取实数 $\lambda > 0$. 令 $X_n^\lambda = -(X_n \wedge \lambda)(n \geqslant 0)$, 则

$$\begin{aligned}\boldsymbol{E}(X_{n+1}^{(\lambda)}|\mathscr{F}_n) &= -\boldsymbol{E}(X_{n+1} \wedge \lambda|\mathscr{F}_n)\\ &\geqslant -(\boldsymbol{E}(X_{n+1}|\mathscr{F}_n) \wedge \boldsymbol{E}(\lambda|\mathscr{F}_n))\\ &= -(X_n \wedge \lambda) = X_n^{(\lambda)} \quad (n \geqslant 0),\end{aligned}$$

即 $X^{(\lambda)} = \{X_n^{(\lambda)}, \mathscr{F}_n, n \geqslant 0\}$ 是下鞅, 显然 $X^{(\lambda)}$ 是一致有界 $\lambda$. 因此, 由 (Ⅱ) 得知

$$\lim_{n\to\infty} Y_n^{(\lambda)} \left(其中\ Y_n^{(\lambda)} = \sum_{k=0}^n V_k D_k^{(\lambda)}, D_k^{(\lambda)} = X_k^{(\lambda)} - X_{k-1}^{(\lambda)}\right)$$

是 a.s. 收敛且有穷的, 所以若令 $X^* = \sup\limits_{n\geqslant 0} |X_n|$, 则

$$\begin{aligned}\{X^* \leqslant \lambda\} &\subset \bigcap_{n\geqslant 0} \{X_n^{(\lambda)} = -X_n\}.\\ &\subset \{\lim_{n\to\infty} Y_n\ 收敛且有穷\}, \quad \text{a.s.},\end{aligned}$$

而由 (1.13)″ 式有 (见定理 1.2)

$$\lim_{\lambda\to\infty} P(X^* \leqslant \lambda) = 1,$$

所以 $\lim\limits_{n\to\infty} Y_n$ 是 a.s. 收敛且有穷的.

(Ⅳ) 设 $X$ 是 $L^1$ 有界鞅, $V^* \leqslant 1$. 推证 $\lim\limits_{n\to\infty} Y_n$ 是 a.s. 收敛且有穷.

由于 $X$ 是 $L^1$ 有界鞅, 所以 $|X| = \{|X_n|, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^1$ 有界非负下鞅. 令

$$W^{(1)} = \{W_n^{(1)}, \mathscr{F}_n, n \geqslant 0\}, \quad W_n^{(1)} = \sup_{k\geqslant n} \boldsymbol{E}(|X_k||\mathscr{F}_n),$$

推证 $W^{(1)}$ 是 $L^1$ 有界非负鞅且 $W_n^{(1)} \geqslant |X_n|$. 事实上, 当 $k \geqslant n$, 时,

$$\begin{aligned}|X_n| \leqslant \boldsymbol{E}(|X_k||\mathscr{F}_n) &\leqslant \boldsymbol{E}(\boldsymbol{E}(|X_{k+1}||\mathscr{F}_k)|\mathscr{F}_n)\\ &= \boldsymbol{E}(|X_{k+1}||\mathscr{F}_n),\end{aligned}$$

所以 $W_n^{(1)} = \lim\limits_{k\to\infty} \boldsymbol{E}(|X_k||\mathscr{F}_n) \geqslant |X_n|(n \geqslant 0)$. 显然 $W_n^{(1)} \in L^1(\Omega, \mathscr{F}, P)(n \geqslant 0)$, 而且

$$\begin{aligned}\boldsymbol{E}(W_{n+1}^{(1)}|\mathscr{F}_n) &= \boldsymbol{E}\left(\lim_{k\to\infty} \boldsymbol{E}(|X_k||\mathscr{F}_{n+1})|\mathscr{F}_n\right)\\ &= \lim_{k\to\infty} \boldsymbol{E}(\boldsymbol{E}(|X_k||\mathscr{F}_{n+1})|\mathscr{F}_n)\\ &= \lim_{k\to\infty} \boldsymbol{E}(|X_k||\mathscr{F}_n)\\ &= W_n^{(1)} \quad (n \geqslant 0).\end{aligned}$$

所以 $\boldsymbol{E}(W_n^{(1)}) = \boldsymbol{E}(W_{n+1}^{(1)})(n \geqslant 0)$,

$$\begin{aligned}\|W^{(1)}\|_1 = \|W_0^{(1)}\| &= \|\sup_{k\geqslant 0}\boldsymbol{E}(|X_k||\mathscr{F}_0)\|_1\\ &= \int_\Omega \lim_{k\to\infty}\boldsymbol{E}(|X_k||\mathscr{F}_0)dP\\ &= \lim_{k\to\infty}\int_\Omega \boldsymbol{E}(|X_k||\mathscr{F}_0)dP = \lim_{k\to\infty}\boldsymbol{E}(|X_k|)\\ &= \sup_{k\geqslant 0}\|X_k\|_1 = \|X\|_1 < \infty.\end{aligned}$$

故 $W^{(1)}$ 是 $L^1$ 有界非负鞅, 且 $\|W^{(1)}\|_1 = \|X\|_1, W_n^{(1)} \geqslant |X_n|$.

再令 $W_n^{(2)} = W_n^{(1)} - X_n(n \geqslant 0)$, 则由 $W^{(1)}$ 和 $X$ 皆为 $L^1$ 有界鞅及 $W_n^{(1)} \geqslant X_n$ 得知 $W^{(2)} = \{W_n^{(2)}, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^1$ 有界非负鞅. 令

$$\begin{aligned}&D_n^{(i)} = W_n^{(i)} - W_{n-1}^{(i)}(i = 1, 2; n = 0, 1, \cdots; W_{-1}^{(i)} \equiv 0),\\ &Y_n^{(i)} = \sum_{k=0}^n V_k D_k^{(i)} \quad (i = 1, 2; n = 0, 1, \cdots),\end{aligned}$$

则

$$Y_n = \sum_{k=0}^n V_k D_k = Y_n^{(1)} - Y_n^{(2)} \quad (n \geqslant 0).$$

而由 (Ⅲ) 得知 $\lim\limits_{n\to\infty} Y_n^{(i)}$ 是 a.s. 收敛且有穷的, 故 $\lim\limits_{n\to\infty} Y_n$ 亦然.

(V) 设 $X$ 是 $L^1$ 有界鞅. 推证 $\lim\limits_{n\to\infty} Y_n$ 在 $\{V^* < \infty\}$ 上 a.s. 收敛且有穷. 令

$$U_k(\lambda) = \begin{cases} V_k, & \text{若 } |V_k| \leqslant \lambda;\\ 0 & \text{反之}\end{cases} \quad (\lambda > 0).$$

则由 (Ⅳ) 得知 $\lim\limits_{n\to\infty}\sum\limits_{k=0}^n \dfrac{U_k(\lambda)}{\lambda} D_k$ 是 a.s. 收敛且有穷的, 故 $\lim\limits_{n\to\infty}\sum\limits_{k=0}^n U_k(\lambda)D_k$ 亦然. 但是对一切 $\lambda > 0$,

$\{V^* \leqslant \lambda\} \subset \bigcap\limits_{n=0}^\infty \{U_n = V_n\} \subset \{\lim\limits_{n\to\infty} Y_n \text{ 收敛且有穷}\}$, a.s., 所以

$$\{V^* < \infty\} \subset \{\lim_{n\to\infty} Y_n \text{ 收敛且有穷}\}, \quad \text{a.s.},$$

定理证毕.

**推论** 设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^1$ 有界鞅, $\{D_n : n \geqslant 0\}$ 是 $X$ 的鞅差序列, 则 $\sum\limits_{n=0}^\infty D_n^2 < \infty$, a.s..

**证**　令 $V_0 \equiv 0, V_n = X_{n-1} (n \geqslant 1)$, 则 $V = \{V_n : n \geqslant 0\}$ 是 $\{\mathscr{F}_n : n \geqslant 0\}$ 可预报序列, $V^* = X^*$, 而 $\|X\|_1 < \infty$, 由定理 1.2 有 $P(V^* < \infty) = P(X^* < \infty) = 1$, 由定理 4.1 得知 $\lim\limits_{n\to\infty} Y_n$ 是 a.s. 收敛且有穷的 $\left(Y_n = \sum\limits_{k=0}^{n} V_k D_k\right)$, 显然 $\lim\limits_{n\to\infty} X_n$ 亦然. 又因为

$$S_n(X)^2 = \sum_{k=0}^{n} D_k^2 = X_n^2 - 2Y_n,$$

所以 $\lim\limits_{n\to\infty} S_n(X)^2 = \sum\limits_{n=0}^{\infty} D_n^2 < \infty$, a.s.. 推论得证.

在这一节以下诸定理中, 若 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅 (上、下鞅), 恒令 $D_n = X_n - X_{n-1} (n \geqslant 0, X_{-1} \equiv 0), S_n(X) = \left(\sum\limits_{k=0}^{n} D_k^2\right)^{\frac{1}{2}} (n \geqslant 0), S(X) = \lim\limits_{n\to\infty} S_n(X) = \left(\sum\limits_{k=0}^{\infty} D_k^2\right)^{\frac{1}{2}}, D_n^* = \sup\limits_{k\leqslant n} |D_k|, D^* = \sup\limits_{n\geqslant 0} |D_n|, X^* = \sup\limits_{n\geqslant 0} |X_n|$.

**定理 4.2**　设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $\boldsymbol{E}(S(X)) < \infty$, 则 $\lim\limits_{n\to\infty} X_n$ 是 a.s. 收敛且有穷的.

**证**　令 $\Omega^* = [0,1], \mathscr{F}^*$ 为 $\Omega^*$ 中一切 Borel 子集. $P^*$ 是 Lebesgue 测度,再令 $\{r_n : n \geqslant 0\}$ 是概率空间 $(\Omega^*, \mathscr{F}^*, P^*)$ 上的相互独立随机变量序列, $r_n$ 只可能取 1 与 $-1$ 两值, 而且 $\int_0^1 r_n(t)dt = 0 (n \geqslant 0)$. 对每个固定的 $t \in [0,1], \left\{\sum\limits_{k=0}^{n} r_k(t) D_k, \mathscr{F}_n, n \geqslant 0\right\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, 从而

$$\left\{\left|\sum_{k=0}^{n} r_k(t) D_k\right|, \mathscr{F}_n, n \geqslant 0\right\}$$

是 $(\Omega, \mathscr{F}, P)$ 上的下鞅, 所以 $\left\{\boldsymbol{E}\left(\left|\sum\limits_{k=0}^{n} r_k(t) D_k\right|\right) : n \geqslant 0\right\}$ 是单调非降序列. 又因为

$$\int_0^1 r_j(t) r_k(t) dt = \int_0^1 r_j(t) dt \int_0^1 r_k(t) dt = 0 \quad (j \neq k),$$

所以

$$\begin{aligned}
\int_0^1 \boldsymbol{E}\left(\left|\sum_{k=0}^{n} r_k(t) D_k\right|\right) dt &= \boldsymbol{E}\left(\int_0^1 \left|\sum_{k=0}^{n} r_k(t) D_k\right| dt\right) \\
&\leqslant \boldsymbol{E}\left(\left[\int_0^1 \left|\sum_{k=0}^{n} r_k(t) D_k\right|^2 dt\right]^{\frac{1}{2}}\right) \\
&= \boldsymbol{E}\left(\left[\int_0^1 \sum_{k=0}^{n} r_k^2(t) D_k^2 dt\right]^{\frac{1}{2}}\right) \\
&= \boldsymbol{E}\left(\left[\int_0^1 \sum_{k=0}^{n} D_k^2 dt\right]^{\frac{1}{2}}\right) = \boldsymbol{E}(S_n(X)).
\end{aligned}$$

再用积分的单调收敛定理 (因为 $\boldsymbol{E}\left(\left|\sum\limits_{k=0}^{n} r_k(t)D_k\right|\right)$ 对 $n$ 单调非降) 可得

$$\int_0^1 \sup_{n\geqslant 0} \boldsymbol{E}\left(\left|\sum_{k=0}^{n} r_k(t)D_k\right|\right) dt \leqslant \boldsymbol{E}(S(X)) < \infty.$$

所以存在 $t_0 \in [0,1]$, 使

$$\sup_{n\geqslant 0} \boldsymbol{E}\left(\left|\sum_{k=0}^{n} r_k(t_0)D_k\right|\right) < \infty,$$

即 $Y = \{Y_n = \sum\limits_{k=0}^{n} r_k(t_0)D_k, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^1$ 有界鞅.

又因为

$$\begin{aligned} X_n &= \sum_{k=0}^{n} D_k = \sum_{k=0}^{n} r_k(t_0)\cdot r_k(t_0)D_k \\ &= \sum_{k=0}^{n} r_k(t_0)(Y_k - Y_{k-1}) \quad (n \geqslant 0), \end{aligned}$$

此即 $X$ 是 $L^1$ 有界鞅 $Y$ 的关于可预报序列 $\{r_k(t_0) : k \geqslant 0\}$ 的变换. 而 $\sup\limits_{n\geqslant 0} |r_n(t_0)| = 1$, 所以由定理 4.1 得知 $\lim\limits_{n\to\infty} X_n$ 是 a.s. 收敛且有穷.

**定理 4.3** 设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}, Y = \{Y_n, \mathscr{F}_n, n \geqslant 0\}$ 皆为概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $X$ 是 $L^1$ 有界的, 且 $S_n(Y) \leqslant S_n(X) (n \geqslant 0)$, 则 $\lim\limits_{n\to\infty} Y_n$ 是 a.s. 收敛且有穷的.

**证** 任取正数 $c > 0$. 令

$$\tau = \inf\{n : |X_n| \geqslant c \text{ 或 } S_n(X) \geqslant c\},$$

并定义 $\inf \varnothing = \infty, S_\infty(X) = S(X)$. 推证

$$\boldsymbol{E}(S_\tau(X)) < \infty.$$

事实上,

$$S_\tau(X) \leqslant \begin{cases} c + |D_\tau| \leqslant 2c + |X_\tau|, & \text{当 } \{\tau < \infty\}, \\ c, & \text{当 } \{\tau = \infty\}, \end{cases} \tag{4.9}$$

若令 $\tau_n = \tau \wedge n$, 则有

$$\begin{aligned} \int_{\{\tau<\infty\}} |X_\tau| dP &\leqslant \lim_{n\to\infty} \inf \int_{\{\tau<\infty\}} |X_{\tau_n}| dP \\ &\leqslant \sup_{n\geqslant 0} \boldsymbol{E}(|X_{\tau_n}|). \end{aligned} \tag{4.10}$$

又因为 $\{|X_n|, \mathscr{F}_n, n \geqslant 0\}$ 是下鞅, $\tau, \tau_n$ 是 $\{\mathscr{F}_k : k \geqslant 0\}$ 停时, 所以

$$\begin{aligned}\boldsymbol{E}(|X_{\tau_n}|) &= \sum_{k=0}^{n} \boldsymbol{E}(|X_k|; \tau_n = k) \\ &\leqslant \sum_{k=0}^{n} \boldsymbol{E}(|X_n|; \tau_n = k) \\ &= \boldsymbol{E}(|X_n|). \end{aligned} \tag{4.11}$$

以 (4.11)、(4.10) 代入 (4.9) 式并注意 $X$ 是 $L^1$ 有界的, 则可得 $\boldsymbol{E}(S_\tau(X)) < \infty$. 令 $\widehat{Y} = \{\widehat{Y}_n = Y_{\tau_n}, \mathscr{F}_{\tau_n} n \geqslant 0\}$, 显然 $\boldsymbol{E}(|\widehat{Y}_n|) < \infty, \widehat{Y}_n \in \mathscr{F}_{\tau_n}$, 且对任何 $A \in \mathscr{F}_{\tau_n}$, 有

$$A \bigcap \{\tau_n = j\} \bigcap \{\tau_{n+1} > j\} \in \mathscr{F}_j,$$

故由 $Y$ 是鞅可得

$$\int_A (Y_{\tau_n} - Y_{\tau_{n+1}}) dP = \sum_{j=0}^{n} \int_{A \cap \{\tau_n = j\} \cap \{\tau_{n+1} > j\}} (Y_j - Y_{j+1}) dP = 0,$$

所以 $\widehat{Y}$ 是鞅. 又因为

$$\begin{aligned}\boldsymbol{E}(S(\widehat{Y})) &= \boldsymbol{E}\left(Y_{\tau_0}^2 + \sum_{n=1}^{\infty} \left|Y_{\tau_n} - Y_{\tau_{n-1}}\right|^2\right)^{\frac{1}{2}} \\ &= \boldsymbol{E}\left(Y_0^2 + \sum_{n=1}^{\tau} |Y_n - Y_{n-1}|^2\right)^{\frac{1}{2}} \\ &= \boldsymbol{E}(S_\tau(Y)) \leqslant \boldsymbol{E}(S_\tau(X)). \end{aligned}$$

所以 $\boldsymbol{E}(S(\widehat{Y})) < \infty$. 因此由定理 4.2 得知 $\lim\limits_{n\to\infty} \widehat{Y}_n$ 是 a.s. 收敛且有穷的. 但是在集合 $\{X^* < c, S(X) < c\}$ 上总有 $\tau = \infty$, 故 $\tau_n = \tau \wedge n = n$, 从而 $Y_n = \widehat{Y}_n (n \geqslant 0)$. 而由定理 1.2 和定理 4.1 系知 $P(X^* < \infty) = 1 = P(S(X) < \infty)$, 所以 $\lim\limits_{n\to\infty} Y_n$ 是 a.s. 收敛且有穷的.

**定理 4.4** 设 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的鞅, $\boldsymbol{E}(D^*) < \infty$, 则

(1) $\lim\limits_{n\to\infty} X_n$ 在 $\{S(X) < \infty\}$ 上是 a.s. 收敛且有穷的;

(2) $\{\sup\limits_{n\geqslant 0} X_n < \infty\} \subset \{S(X) < \infty\}$, a.s..

更一般地, 若还有 $Y = \{Y_n, \mathscr{F}_n, n \geqslant 0\}$ 是 $X$ 关于可预报序列 $V = \{V_n : n \geqslant 0\}$ 的变换, 则

(3) $\lim\limits_{n\to\infty} Y_n$ 在 $\{S(Y) < \infty\} \bigcap \{V^* < \infty\}$ 上是 a.s. 收敛且有穷的;

(4) $\{\sup\limits_{n\geqslant 0} Y_n < \infty\} \bigcap \{V^* < \infty\} \subset \{S(Y) < \infty\}$, a.s.

**注** (a) 设 $\boldsymbol{E}(D^*)<\infty$. 由定理 3.4 及定理 4.4 的 (1) 和 (2) 得 $\{\lim_{n\to\infty} X_n$ 收敛且有穷$\}=\{\sup X_n<\infty\}=\{S(X)<\infty\}$, a.s..

(b) 设 $\boldsymbol{E}(D^*)<\infty, V^*<K$,a.s., 则

$$\boldsymbol{E}\left(\sup_{n\geqslant 0}|Y_n-Y_{n-1}|\right)=\boldsymbol{E}\left(\sup_{n\geqslant 0}|V_nD_n|\right)\leqslant K\boldsymbol{E}(D^*)<\infty,$$

故由定理 3.4 及定理 4.4 的 (iii), (iv) 得

$$\begin{aligned}\{\lim_{n\to\infty} Y_n \text{ 收敛且有穷}\}&=\{\sup Y_n<\infty\}\\&=\{S(Y)<\infty,\text{a.s.}\}.\end{aligned}$$

**证** 设鞅 $X$ 满足 $\boldsymbol{E}(D^*)<\infty$.

(1) 任取正数 $\lambda$. 令 $\tau=\inf\{n: S_n(X)\geqslant\lambda\}, \tau_n=\tau\wedge n, \widehat{X}_n=X_{\tau_n}(n\geqslant 0)$, 则仿定理 4.3 可证 $\widehat{X}=\{\widehat{X}_n,\mathscr{F}_{\tau_n},n\geqslant 0\}$ 是鞅, 显然

$$\begin{aligned}\boldsymbol{E}(S(\widehat{X}))&=\boldsymbol{E}\left(\widehat{X}_0^2+\sum_{k=1}^{\infty}\left|\widehat{X}_k-\widehat{X}_{k-1}\right|^2\right)^{\frac{1}{2}}\\&=\boldsymbol{E}\left(X_{\tau_0}^2+\sum_{k=1}^{\infty}\left|X_{\tau_k}-X_{\tau_{k-1}}\right|^2\right)^{\frac{1}{2}}\\&=\boldsymbol{E}\left(X_0^2+\sum_{k=1}^{\tau}|X_k-X_{k-1}|^2\right)^{\frac{1}{2}}\\&=\boldsymbol{E}(S_{\tau-1}(X)^2+|D_\tau|^2)^{\frac{1}{2}}\\&\leqslant\boldsymbol{E}(\lambda^2+D^{*2})^{\frac{1}{2}}\leqslant\lambda+\boldsymbol{E}(D^*),\end{aligned}$$

所以 $\boldsymbol{E}(S(\widehat{X}))<\infty$. 因此, 由定理 4.2 得 $\lim_{n\to\infty}\widehat{X}_n$ 是 a.s. 收敛且有穷的. 但在集合 $\{S(X)<\lambda\}$ 上 $\tau=\infty$ 且 $X\equiv\widehat{X}$, 所以 $\lim_{n\to\infty}X_n$ 在 $\{S(X)<\lambda\}$ 上是 a.s. 收敛且有穷的. 由于 $\lambda>0$ 可以任意取, 所以 (1) 获证.

(2) 任取 $\lambda>0$, 令 $\tau=\inf\{n:|X_n|\geqslant\lambda\}, \tau_n=\tau\wedge n, \widehat{X}_n=X_{\tau_n}(n\geqslant 0)$. 则 $\widehat{X}=\{\widehat{X}_n,\mathscr{F}_{\tau_n},n\geqslant 0\}$ 是鞅, 显然, $|\widehat{X}_n|\leqslant\lambda+D^*$, 所以 $\widehat{X}$ 是 $L^1$ 有界鞅, 由定理 4.1 系得

$$S(\widehat{X})<\infty,\text{a.e..}$$

但是在集合 $\{X^*<\lambda\}$ 上有 $\tau=\infty$ 且 $S(X)=S(\widehat{X})$, 所以对一切 $\lambda>0$, 有

$$\{X^*<\lambda\}\subset\{S(X)<\infty\},\quad\text{a.s.},$$

从而

$$\{X^*<\infty\}\subset\{S(X)<\infty\},\quad\text{a.s.},$$

但是, 由定理 3.4 有

$$\{X^* < \infty\} = \{\sup_{n\geqslant 0} X_n < \infty\} = \{\inf_{n\geqslant 0} X_n > -\infty\}, \quad \text{a.s..}$$

(2) 获证.

(3) 再设 $Y$ 是 $X$ 的关于可预报序列 $V$ 的变换. 任取 $\lambda > 0$, 令

$$\widehat{V}_n(\omega) = \begin{cases} V_n(\omega), & \text{当 } |V_n(\omega)| < \lambda, \\ 0, & \text{反之} \end{cases} \quad (n \geqslant 0),$$

$$\widehat{Y}_n = \sum_{k=0}^{n} \widehat{V}_k D_k \quad (n \geqslant 0),$$

$$\widehat{G}_n = \widehat{Y}_n - \widehat{Y}_{n-1} \quad (n \geqslant 0, \widehat{Y}_{-1} \equiv 0),$$

$$\widehat{G}^* = \sup_{n\geqslant 0} |\widehat{G}_n|,$$

则 $\widehat{Y} = \{\widehat{Y}_n, \mathscr{F}_n, n \geqslant 0\}$ 是鞅, 且 $\boldsymbol{E}(\widehat{G}^*) \leqslant \lambda \boldsymbol{E}(D^*) < \infty$. 因此, 由 (1) 得知 $\lim\limits_{n\to\infty} \widehat{Y}_n$ 在 $\{S(\widehat{Y}) < \infty\}$ 上是 a.s. 收敛且有穷的. 由于

$$\{V^* < \lambda\} \subset \bigcap_{n=0}^{\infty} \{\widehat{V}_n = V_n\} \subset \bigcap_{n=0}^{\infty} \{\widehat{Y}_n = Y_n\},$$

$$S(\widehat{Y})^2 = \sum_{k=0}^{\infty} (\widehat{V}_k D_k)^2 \leqslant \sum_{k=0}^{\infty} (V_k D_k)^2 = S(Y)^2,$$

所以

$$\begin{aligned} \{V^* < \lambda\}\{S(Y) < \infty\} &\subset \{S(\widehat{Y}) < \infty\} \bigcap_{n\geqslant 0} \{\widehat{Y}_n = Y_n\} \\ &\subset \left\{\lim_{n\to\infty} \widehat{Y}_n \text{ 收敛且有穷}\right\} \bigcap_{n\geqslant 0} \{\widehat{Y}_n = Y_n\} \\ &\subset \left\{\lim_{n\to\infty} Y_n \text{ 收敛且有穷}\right\}, \quad \text{a.} \end{aligned}$$

即 (3) 成立.

(4) 由于 $\boldsymbol{E}(\widehat{G}^*) < \infty$, 对 $\widehat{Y}$ 用 (2) 得

$$\left\{\sup_{n\geqslant 0} \widehat{Y}_n < \infty\right\} \subset \left\{S(\widehat{Y}) < \infty\right\}, \quad \text{a.s.,}$$

再注意

$$\{V^* < \lambda\} \subset \bigcap_{n\geqslant 0} \{\widehat{Y}_n = Y_n\},$$

则可得

$$\left\{\sup_{n\geqslant 0} Y_n < \infty, V^* < \lambda\right\} = \left\{\sup_{n\geqslant 0} \widehat{Y}_n < \infty, V^* < \lambda\right\}$$
$$\subset \{S(Y) = S(\widehat{Y}) < \infty\}, \quad \text{a.s..}$$

由于 $\lambda > 0$ 可以任意选取, 在上式中令 $\lambda \to \infty$, 得

$$\left\{\sup_{n\geqslant 0} Y_n < \infty, V^* < \infty\right\} \subset \{S(Y) < \infty\}, \quad \text{a.s..}$$

**定理 4.5** 设 $(\Omega, \mathscr{F}, P)$ 为概率空间, $A \in \mathscr{F}$ 称 $A$ 是原子集合, 如果 $P(A) > 0$, 但 $A$ 不能包含两个不交的具有正概率的子集. 若 $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 是 $L^1$ 有界鞅, $\{D_n : n \geqslant 0\}$ 是 $X$ 的鞅差序列, $A$ 是原子集合, $D_n$ 的值域皆为可数集, 则

$$\sum_{n=0}^{\infty} |D_n| < \infty \quad, \text{a.s. in } A.$$

**证** 令 $D_n$ 之值域皆含于 $\{a_1, a_2, a_3, \cdots\}, n \geqslant 0$. 由 $A$ 是原子集合知: 固定任一非负整数 $n$, 必存在实数 $a_{k_n}$ 使 $A \subset \{D_n = a_{k_n}\}$, a.s., 亦即 $P(A - \{D_n = a_{k_n}\}) = 0$, 如果不然, 则对一切 $a_k$, 均有 $P(A - \{D_n = a_k\}) > 0$. 令

$$K = \sup\{k : P(A, D_n \neq a_1, \cdots, D_n \neq a_k) = P(A) > 0, k \geqslant 1\},$$

则 $1 \leqslant K \leqslant \infty$.

(a) 若 $K < \infty$, 则

$$P(A, D_n \neq a_1, \cdots, D_n \neq a_k, D_n = a_{k+1}) > 0.$$

因此 $A \bigcap \{D_n \neq a_1, \cdots, D_n \neq a_k, D_n = a_{k+1}\}$ 与 $A \bigcap \{D_n \neq a_{k+1}\}$ 是 $A$ 中 2 个不交的具有正概率的子集, 这与 $A$ 是原子集合矛盾.

(b) 若 $K = \infty$, 则

$$P(A, D_n \neq a_1, D_n \neq a_2, \cdots) = P(A) > 0,$$

这与 $P\left(\bigcup_{k=1}^{\infty} \{D_n = a_k\}\right) = 1$ 矛盾. 所以必存在 $a_{k_n}$ 使

$$A \subset \{D_n = a_{k_n}\}, \quad \text{a.s.,}$$

从而 $A \subset \bigcap_{n=0}^{\infty} \{D_n = a_{k_n}\}$,[a.e.]. 取

$$V_n = \begin{cases} 1, & \text{若 } a_{k_n} \geqslant 0, \\ -1, & \text{若 } a_{k_n} < 0, \end{cases}$$

则 $X$ 关于 $V$ 的变换 $Y = \left\{\sum_{k=0}^{n} V_k D_k, \mathscr{F}_n, n \geqslant 0\right\}$ 满足

$$Y_n = \sum_{j=0}^{n} |a_{k_j}| = \sum_{j=0}^{n} |D_j|, \quad \text{a.s. in } A.$$

再利用定理 4.1 可证定理 4.5.

## §5　习题及应用

1. 证明引理 2.1.
2. 证明引理 2.2.
3. 证明引理 2.3.
4. 证明命题 3.1.
5. 证明命题 3.2.
6. 证明命题 3.3.
7. 证明命题 3.4.
8. 证明命题 3.5.
9. 试构造一个随机过程, 它是 Markov 过程而不是鞅.
10. 试构造一个随机过程, 它既是 Markov 过程又是鞅.
11. 试构造一个随机过程, 它是鞅而不是 Markov 过程.

12. (条件期望的连续性) 设 $(\Omega, \mathscr{F}, P)$ 是任一概率空间, $\{\mathscr{F}_n, n = 0, 1, \cdots\}$ 是 $\mathscr{F}$ 中的一族子 $\sigma$ 代数, $\xi \in L^1(\Omega, \mathscr{F}, P)$. 试证:

(1) 若 $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \cdots$, 则

$$\lim_{n\to\infty} \boldsymbol{E}(\xi|\mathscr{F}_n) = \boldsymbol{E}\left(\xi \Big| \bigvee_{n\geqslant 0} \mathscr{F}_n\right), \quad \text{a.s. 及 } L^1.;$$

(2) 若 $\mathscr{F}_0 \supset \mathscr{F}_1 \supset \mathscr{F}_2 \supset \cdots$, 则

$$\lim_{n\to\infty} \boldsymbol{E}(\xi|\mathscr{F}_n) = \boldsymbol{E}\left(\xi \Big| \bigcap_{n\geqslant 0} \mathscr{F}_n\right), \quad \text{a.s. 及 } L^1..$$

13. 设 $(\Omega, \mathscr{F}, P)$ 是任一概率空间, $\{\mathscr{F}_n, n = 0, 1, \cdots\}$ 是 $\mathscr{F}$ 中一族单增的子 $\sigma$ 代数且 $\mathscr{F} = \bigvee_{n\geqslant 0} \mathscr{F}_n$, 则

(1) $\bigcup_{n\geqslant 0} L^1(\Omega, \mathscr{F}_n, P)$ 在 $L^1(\Omega, \mathscr{F}, P)$ 中稠密, 即是对任何 $\xi \in L^1(\Omega, \mathscr{F}, P)$, 存在 $\{\xi_n, n \geqslant 0\} \subset \bigcup_{n\geqslant 0} L^1(\Omega, \mathscr{F}_n, P)$, 使

$$\lim_{n\to\infty} \xi_n = \xi, \quad L^1.$$

(2) 对任何 $\xi \in L^1(\Omega, \mathscr{F}, P)$, 存在 $\{\xi_n, n \geqslant 0\} \subset \bigcup\limits_{n \geqslant 0} L^1(\Omega, \mathscr{F}_n, P)$, 使

$$\lim_{n\to\infty} \xi_n = \xi, \quad \text{a.s.}.$$

14. 设 $\Omega = [0,1)$, $\mathscr{F}$ 为 $\Omega$ 中的全体 Borel 集, $P$ 是 $\mathscr{F}$ 上的 Lebesgue 测度. 试构造 $\mathscr{F}$ 中的一族单增子 $\sigma$ 代数 $\{\mathscr{F}_n, n = 0, 1, \cdots\}$, 使 $\mathscr{F} = \bigvee\limits_{n \geqslant 0} \mathscr{F}_n$, 且对任何 $\xi \in L^1(\Omega, \mathscr{F}, P)$, 存在 $\{\xi_k, k \geqslant 0\} \subset \bigcup\limits_{n \geqslant 0} L^1(\Omega, \mathscr{F}_n, P)$, $\{\eta_k, k \geqslant 0\} \subset \bigcup\limits_{n \geqslant 0} L^1(\Omega, \mathscr{F}_n, P)$, 使

$$\xi = \lim_{k\to\infty} \xi_k, L^1.;$$
$$\xi = \lim_{k\to\infty} \eta_k, \text{a.s.}.$$

15. 设 $L^1(\Omega, \mathscr{F}, P)$ 如习题 14. 试证: $[0,1)$ 上的全体连续函数在 $L^1(\Omega, \mathscr{F}, P)$ 中稠, 而且对任何 $\xi \in L^1(\Omega, \mathscr{F}, P)$, 还存在 $[0,1)$ 上的一族连续函数 $\{\eta_n, n \geqslant 0\}$ 使

$$\lim_{n\to\infty} \eta_n = \xi, \quad \text{a.s.}.$$

16. 设 $L^1(\Omega, \mathscr{F}, P)$ 如习题 14. 试证: 任取 $\xi \in L^1(\Omega, \mathscr{F}, P)$, 有

$$\lim_{n\to\infty} \frac{1}{2^n} \sum_{k=0}^{\frac{1}{2^n}-1} \xi\left(x + \frac{k}{2^n}\right) = \int_0^1 \xi(t)dt, \quad \text{a.s.},$$

其中 $\xi$ 是以 1 为周期的周期函数.

17. 设可测空间 $(\Omega, \mathscr{F})$ 是可分的, 即存在 $\{A_n : n \geqslant 0\} \subset \mathscr{F}$, $\sigma(\{A_n, n \geqslant 0\}) = \mathscr{F}$. 令 $\mathscr{F}_n = \sigma(A_0, A_1, \cdots, A_n)$, $\pi_n = \{B_1^{(n)}, \cdots, B_{k_n}^{(n)}\}$, $B_i^{(n)} \bigcap B_j^{(n)} = \varnothing$, $\bigcup\limits_{i=1}^{k_n} B_i^{(n)} = \Omega$, $\pi_n < \pi_{n+1}$ (即 $\pi_n$ 中任一集均可表示为 $\pi_{n+1}$ 中有限个集合之并), $\sigma(\pi_n) = \mathscr{F}_n (n \geqslant 0)$, $P$ 是 $\mathscr{F}$ 上的概率测度, $\varphi$ 是 $\mathscr{F}$ 上的有限测度, $\varphi \ll P$ (即 $P(A) = 0 \Rightarrow \varphi(A) = 0$), $\xi = \dfrac{d\varphi}{dP}$ 为 $\varphi$ 对 $P$ 的 Radon–Nikodym 导数. 再令

$$X_n(\omega) = \sum_{A \in \pi_n} \frac{\varphi(A)}{P(A)} \mathbf{1}_A(\omega) \quad (n \geqslant 0, \omega \in \Omega),$$

约定 $\dfrac{0}{0} = 0$, 试证: $X = \{X_n, \mathscr{F}_n, n \geqslant 0\}$ 为一致可积鞅, 且

$$\lim_{n\to\infty} X_n = \xi = \frac{d\varphi}{dP}, \quad \text{a.s.}.$$

18. (Kolmogorov 零一律) 设 $\{X_n, n \geqslant 0\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一列随机变量, $T_n \overset{\text{def.}}{=\!=} \sigma(X_{n+1}, X_{n+2}, \cdots)$, $T_\infty = \bigcap\limits_{n=0}^{\infty} T_n$. 称 $T_\infty$ 为尾 $\sigma$ 代数, $T_\infty$ 中

每个集合 $A$ 皆称为 $\{X_n, n \geqslant 0\}$ 的一个尾集合. 若 $\{X_n, n \geqslant 0\}$ 相互独立, 试证: $\{X_n, n \geqslant 0\}$ 的每一个尾集合 $A$, 均有:

$$P(A) = 0 \quad \text{或} \quad 1.$$

19. 设 $\{X_n, n \geqslant 0\}$ 如习题 18, $\{a_n, n \geqslant 0\}$ 为一列实数, 则

$$P\left(\sum_{n=0}^{\infty} a_n X_n \text{ 收敛且有穷}\right) = 0 \text{ 或 } 1.$$

20. 设 $(\Omega, \mathscr{F}, P), \{X_n, n \geqslant 0\}, \{T_n, n \geqslant 0\}$ 如习题 18. 若 $\{X_n, n \geqslant 0\}$ 还有公共分布, 试证:

(1) $\boldsymbol{E}(|X_n|) < \infty \Rightarrow \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k \to \boldsymbol{E}(X_0)$ a.s.;

(2) $X_n \geqslant 0, \boldsymbol{E}(X_n) = \infty \Rightarrow \dfrac{1}{n}\sum\limits_{k=1}^{n} X_k \to \infty$, a.s..

21. 试用鞅论证明: 若 $\{X_t, t \in [0, \infty)\}$ 是概率空间 $(\Omega, \mathscr{F}, P)$ 上的一个一维的标准 Brown 运动, 则对任何实数 $a$ 及正实数 $\lambda > 0$, 恒有

$$\lambda P\left(\sup_{t \geqslant 0} e^{aX_t - \frac{ta^2}{2}} > \lambda\right) \leqslant 1.$$

# 参 考 文 献


[1] Ash R B, Garduer M F. Topics in stochastic processes. New York: Academic Press, 1975.

[2] Athery K B, Karlin S. On branching processes with random environment, I. Extinction probabilities. Ann. Math. Siatist, 42, 1971, 1499-1520.

[3] Bertion J. Levy processes. Cambridge: Cambridge university Press. 1996.

[4] Billingsley P. Convergence of probability measure. New York: John Wiley and Sons, 1968.

[5] Blumenthal R M, Getoor R K. Markov processes and potential theory. New York: Academic Press, 1968.

[6] Burkholder D L. Martingale transforms. Ann. Math. Statist., 37, 1966: 1494-1504.

[7] Brukholder D L. A sharp inequality for matringale transforms. Ann. Probab., 7, 1979, 858-863.

[8] Burkholder D L. A geometrical characterization of Banach space in which martingale difference sequences are unconditional. Ann. Probab., 9, 1987, 997-1011.

[9] Chacon R V, Ornstein D S. A general ergodic theorem. Illinois J. Math., 4, 1960, 153-160.

[10] Chatterji S D. Martingale convergence and the Radon-Nikodym theorem in Banach space. Math. Scand., 22, 1968, 21-41.

[11] 钱敏, 侯振挺等. 可逆马尔可夫过程. 长沙: 湖南科学技术出版社, 1979.

[12] 陈木法, 郑小谷. q 过程的惟一性准则. 中国科学 (A). 12, 1982, 288-308.

[13] Chung K L. Markov Chains with Stationary Transition Probabilities, 2nd. New York: Springer, 1967.

[14] Cogbum R. Markov chains in random environments. The case of Markovian environments. Ann. Probab., 8, 1980, 908-916.

[15] Cogburn R. The ergodic theory of Markov chains in random environments. Z. Wahrsch. Verw. Gebiete, 66, 1984, 109-128.

[16] Cogburn R. On direct convergence and periodicity for transition probabilities of Markov chains in random environments. Ann. Probab., 18, 1990, 642-654.

[17] Cogburn R. On the central limit theorem for Markov chains in random environments. Ann. Probab., 19, 1991, 587-604.

[18] Cohn H. A martingale approach to supercritical branching processes. Ann. Probab., 13, 1985, 1179-1191.

[19] Cohn H. On the growth of the multitype to supercritical branching processes in a random environment. Ann. Probab., 17, 1989,1118-1123.

[20] Donsker M. An invariance principle for certain probability limit theorems, Memoirs of Amer. Math. Soc., 6, 1951.

[21] Dellacheric C. Capacites et processus stochastiques. New York: Springer-Verlag, 1972.

[22] Diestel J, Uhl J J Jr. Vector measure. Mathematical Surveys and Monographs, 15, AMS, Providence, Rhode Island, 1977.

[23] Doob J L. Stochastic Processes. John Wiley and Sons, 1953.

[24] Дынкин Е Б. Оснований Теории Марковских процессов. Москва: Государставеиное Издателъство Фиэико—Математическон Литературы, 1959. (Denkin E B. Basic theory of Markov processes. Moscow: Mathematical and Physical Press, 1959. (in Russian))

[25] Дынкин Е Б. Марковские Процессы. Москва: Государставеиное И эдателъство Фиэико—Матиматнческой Литсратуры,1963.
(Denkin E B. Markov processes. Moscow: Mathematical and Physical Press, 1963. (in Russian))

[26] Дыкин Е Б, Ющкевич А А. Строго марковские процессы Теор. и её прим., 1, 1956, 149-156.
(Denkin E B, Ushkaivich A A. Strong Markov processes, Probability Theory and its Applications, 1, 1956, 149-156. (in Russian))

[27] Federer H. Geometric measure theory. New York: Springer-Verlag, 1969.

[28] Feller W. An introduction to probability theory and its applications(Vol. 2) Second edition. New York: John Wiley and Sons, 1971.

[29] Getoor P K. Markov processes, Ray processes and right processes. Lecture Notes in Math., 440, Springer-Verlag, 1975.

[30] Гихман И И, Скороход А В. Введение в теорию случаиных прочессов иэдательство Москва: Наука, 1965.

[31] Halmos P R. Measure theory. New York: D. Van Nostrand Company INC., 1950.

[32] Harris T E. The theory of branching processes. New York: Springer-Verlag, 1963.

[33] Hille E. Functional analysis and semi-groups. Amer. Math. Soc. Colloquium Publications, 31, 1948.

[34] 侯振挺, 郭青峰. 齐次可列马尔可夫过程. 北京: 科学出版社, 1978.

[35] 胡迪鹤. 随机过程论 —— 基础, 理论, 应用 (第二版). 武汉: 武汉大学出版社, 2005.

[36] 胡迪鹤. 分析概率论 (第二版). 北京: 科学出版社, 1997.

[37] 胡迪鹤. 一般状态马氏过程分析理论. 武汉: 湖北教育出版社, 1985.

[38] 胡迪鹤. 可数状态的马尔可夫过程论. 武汉: 武汉大学出版社, 1983.

[39] 胡迪鹤, 甘师信. 近代鞅论. 武汉: 武汉大学出版社, 1993.

[40] 胡迪鹤等. 随机分形引论. 武汉: 武汉大学出版社, 1996.

[41] 胡迪鹤, 不变原理及其在分支过程中的应用. 北京大学学报, 10(1), 1964, 1-26.

[42] 胡迪鹤. 马尔可夫链的泛函的极限分布. 武汉大学学报, 23(3), 1977, 63-79.

[43] 胡迪鹤. 抽象空间中 q- 过程的构造理论. 数学学报, 16(2), 1996, 150-165.

[44] 胡迪鹤. 度量空间中转移函数的强连续性, Feller 性及强马尔可夫性. 数学学报, 20(4), 1997, 298-300.

[45] 胡迪鹤. 抽象空间中 q- 过程的构造理论 (Ⅱ). 数学学报, 21(2), 1978, 190-192.

[46] 胡迪鹤. 非时齐的可数状态的 Q- 过程的存在惟一性. 数学学报, 22(3), 1978, 285-287.

[47] 胡迪鹤. 关于某些随机阵的调和函数. 数学学报, 22(3), 1979, 276-290.

[48] 胡迪鹤. 抽象空间中非时齐马氏过程的分析理论 (I). 数学学报, 22(4), 1979, 420-437.

[49] 胡迪鹤. 抽象空间中非时齐马氏过程的分析理论 (Ⅱ). 数学学报, 22(5), 1979, 530-545.

[50] 胡迪鹤. 抽象空间中非时齐马氏过程的分析理论 (Ⅲ). 数学学报, 22(6), 1979, 643-651.

[51] 胡迪鹤. 抽象空间中 q- 过程的惟一性准则. 数学学报, 23(5), 1980, 750-757.

[52] 胡迪鹤. 关于 B 值鞅及经典分析. 应用概率统计, 2(4), 1986, 362-370.

[53] 胡迪鹤. Hillbert 空间上各向同性的马氏场 (I). 数学杂志, 3(1), 1983, 35-54.

[54] 胡迪鹤. Hillbert 空间上各向同性的马氏场 (II). 数学杂志, 3(1), 1983, 145-156.

[55] 胡迪鹤. 抽象空间中马氏过程的强遍历性及收敛速度. 数学学报, 27(3), 1984, 293-304.

[56] 胡迪鹤. 抽象空间中 q- 过程遍历位势. 数学学报, 27(4), 1984, 469-481.

[57] 胡迪鹤. 马氏场的遍历性及耦合. 数学学报, 35(4)1992, 505-515.

[58] Hu Dihe. The necessary and sufficient conditions of self-similar sets snd their dimension. Stochastic Processes and Their Applications, 90, 2000, 243-262.

[59] Hu Dihe. Probability properties and fractal properites of statistically recursive sets. Science in China (Series A), 44, 2001, 742-761.

[60] Hu Dihe. The construction of Markov processes in random environments and the equivalence theorems. Science in China (Series A), 47, 2004, 481-496.

[61] Hu Dihe. The existence and uniqueness of q-process in random environments. Science in China (Series A), 47, 2004, 641-658.

[62] Hu Dihe. Infinitely dimensional control, Markov brinching chains in random environments. Science in China (Series A), 49(1), 2006, 27-53.

[63] Hu Dihe. I. I. D. ststistical contration operators and statistically self-similar sets. Chin. Ann. Math., 23B, 2002, 461-468.

[64] Hu Dihe. The structure and approximation of A, S. self-similar sets. Acta Math. Sci. 23B, 2003, 201-207.

[65] Hu Dihe. The existence and moments of cononical branching chain in random environment. Acta Math. Sci., 24B3, 2004, 499-506.

[66] Hu Dihe. The classification. and period of states for Markov chains in random environments. Acta Math. Sci., 25B1, 2005, 13-29.

[67] Hu Dihe. The decompositon of state space for Markov chains in random environments. Acta Math. Sci., 25B, 2005, 555-568.

[68] Hu Dihe. From p-m chains to Maekov chains in random environments. Chin. Ann. Math., 26A, 2004, 65-78.

[69] Hu Xiaoyu, Hu Dihe. The dimensions of the zero sets of a recurrent random walk. Chinese J. of Contemparary Maht., 16(3), 1995, 275-282.

[70] Hu Dihe, et. al. Summary of Recent Research Accomplishments in Markov Processes and Markov Fields at Wuhan University. Contemporay Math. (U. S. A. ), 118, 1191, 149-168.

[71] Hu Dihe. The relation among various Markov chains. Wuhan Univ. J. 6, 2001, 543-548.

[72] Hu Dihe Xiao Z. The invariance principle for p−θ chain. Acta Math. Sinica, (English Series), 23, 2007, 41-56.

[73] Hu Dihe. The constuction of multitype canonical Markov branching chains in random environments. Aata Math. Sci., 26B, 2006, 431-442.

[74] Hu Dihe, Hu Xiaoyu. The construction of denumerable q-processes in random environments. Acta Math. Sci., 28B, 2008, 225-235.

[75] Hu Dihe. Double conditional expectation. Wuhan Uni. J. 9(6), 2004, 851-857.

[76] Hu Dihe, Zhang Shulin. The Laplace functionals and moments for Markov branching chains in random environments. Wuhan Uni. J., 10, 2005, 485-492.

[77] Hu Dihe, Hu Xiaoyu. The equivalence forms of random Kolmogorov forwarld (backward) equitions. Wuhan Uni. J., 10, 2005, 808-812.

[78] 胡迪鹤等. 随机矩阵的随机调和函数. 武汉大学学报, 51, 2005, 1-6.

[79] 胡迪鹤. 随机坏境中的生灭过程的基本概念及存在性. 武汉大学学报, 49, 2003, 1-5.

[80] Ibragimov I A, Rozanov Y A. Gaussian random processes. Translared by Aries A. B. New York: Springer-Verlag, 1978.

[81] 伊藤清. 随机过程. 刘璋温译. 上海: 上海科学技术出版社, 1961.

[82] Ito K, Mckean H P Jr. Diffusion processes and their sample path. New York: Academic Press, 1965.

[83] Kalinkin A V. Markov branching processes with interaction. Russian Math Surveys, 244-304.

[84] Kelley J L. General topology. New York:Spinger-Verlag, 1955.

[85] Kendall D G. Some analytical properties of continuos stationary Markov transition functions. Tran. Amer. Math. Soc., 78, 1955, 529-540.

[86] Kifer Y. Limit theorems for random transformations and processes in random environments. Tran. Amer. Math. Soci., 350, 1998, 1481-1518.

[87] Kingman J F C. The ergodic theory of subadditve stochastic processes. J. Royal Stat. Soc. B, 30, 499-520.

[88] Kingman J F C. Subadditive proesses. Lecture Notes Math. 539, New York: Springer-Verlag, 1976.

[89] Li Y. The recurrence and invariant measures for Markov chains in bi-infnite environments, Science in China (Series A), 31, 2001, 702-707.

[90] Loeve M. Probability theory($4^{th}$ edition). New York: Springer-Verlag, 1977.

[91] Meyer P A. Probability and potential. Hermann, 1996.

[92] Meyer P A, Processus de Markov. Lecture Notes Math. 26, New York: Springer-Verlag, 1967.

[93] Nawrotzki K., Discrete open system on Markov chains in a random envionment, I, II. J. Inform. Process, Cgbernet, 17, 1981, 569-599; 18, 1982, 83-98.

[94] Neveu J. Discrete parameter martingales. Translated by Speed T P. Amsterdam: North Hollard Publishing Company, 1976.

[95] Orey S. Markov chains wish stochastically stationary transition probabilities. Ann. Probab., 19, 1991, 907-928.

[96] Parzen E. Stochastic processes. Holden-Day, Inc., 1962.

[97] Ray D. Resolvents, Transitions and Strongly Msrkovian processes. Ann. Math., 70, 1959, 43-72.

[98] Reuter G E H. Denumerable Markov processes and the associated contraction semigroup. L. Acta Math., 97, 1957, 1-46.

[99] Reuter G E H. Denumerable Markov processes (II). (III). J. London Math. Soc., 34, 1959, 81-91; 37, 1962, 63-73.

[100] Riesz F, Nagy B Sz. Functional Analysis. 1956.

[101] Richart C E. General theory of Banach algebras. New York: Robert E. Krieger Publishing Co. Inc., 1974.

[102] Rockafeller R T. Convex analysis. New York: Princeton University Press, 1972.

[103] Rozanov Yu A. Stationary random processes. SanFrancisco: Holden-day, 1967.

[104] Tanny D, On multitype branching processes in a random environmrnt. Adv. Appl. Probab., 13, 1981, 464-49.

[105] Taylor R L. Stochastic convergence of weighted sums of random elements in linear space. Lecture Notes in Math., 672, Springer-Verlag, 1978.

[106] Walters P, An introducion to ergodic theory. New York: Springer-Verlag, 1981.

[107] 王梓坤. 生灭过程与马尔可夫链. 北京: 科学出版社, 1980.

[108] 王梓坤. 随机过程论. 北京: 科学出版社, 1978.

[109] 王梓坤, 杨向群. The birth and death processes and Markov chains. New York: Springer-Verlag, 1992.

[110] Wider D V. The Laplace transform. New York: Princeton University Press, 1946.

[111] 吴智泉, 王向忱. 巴氏空间上的概率论. 长春: 吉林大学出版社, 1990.

[112] 肖争艳, 胡迪鹤. 绕积马氏链的状态分类. 数学物理学报, 23A(3) 2003, 306-313,

[113] 薛乃华, 肖争艳, 胡迪鹤. 绕积 Markov 链的不变测度及遍历极限. 武汉大学学报, 47, 2001, 17-21.

[114] 杨向群. 可列马尔科夫过程构造论. 长沙: 湖南科学技术出版社, 1980.

[115] 严士健, 王隽骧, 刘季芳. 概率论基础. 北京: 科学出版社, 1982.

[116] 严家安. 鞅与随机积分引论. 上海: 上海科学技术出版社, 1981.

[117] Yosida K. Functional analysis (sixth edition). New York: Springer-Verlag, 1980.

[118] 施仁杰, 卢科学. 时间序列分析引论. 西安: 西安电子科技大学出版社, 1988.

[119] 肖争艳, 胡迪鹤. 随机环境中多维分枝链的增长率及灭绝概率. 数学进展, 35, 2006, 685-698.

# 索 引

## B



## C



## D



## E



## F

## R



## S



## T



## W



## X



## Y

## Z